

PROCEEDINGS



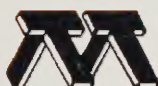
1998 IEEE Radio and Wireless Conference

Sheraton Colorado Springs

Colorado Springs, Colorado, USA

August 9-12, 1998

<http://rawcon.org>



NIST ITS

RAWCON'98 Co-Sponsors

IEEE Microwave Theory and Techniques Society

IEEE Pikes Peak Section

RAWCON'98 Technical Co-Sponsors

IEEE Communications Society

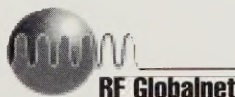
National Institute of Standards and Technology

Institute for Telecommunications Sciences

RAWCON'98 Media Co-Sponsors

Applied Microwave & Wireless

RF Globalnet Website

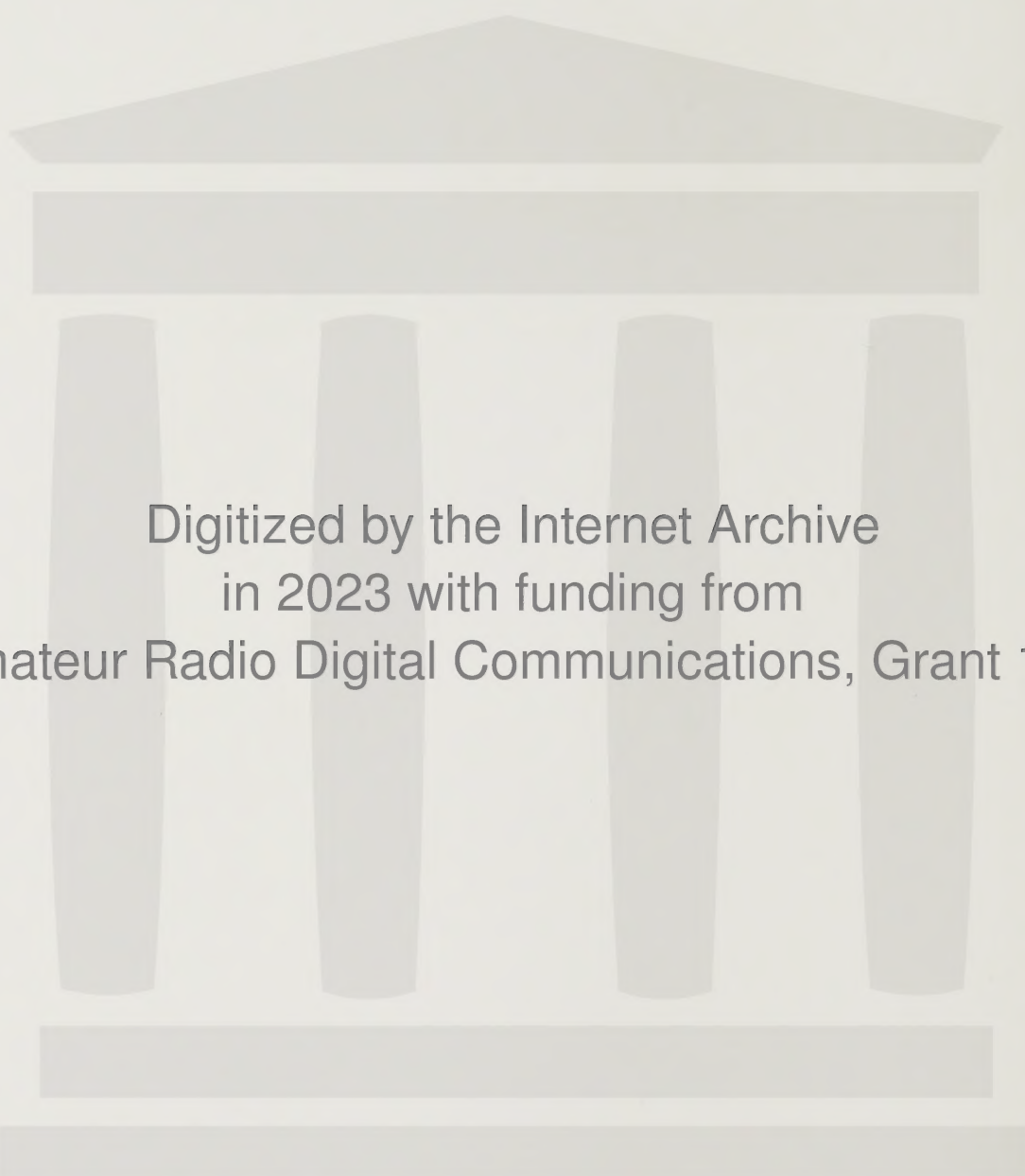


General Chair:

Dr. Roger Marks, *National Institute of Standards and Technology*

Technical Program Chair:

Dr. Michael S. Heutmaker, *Lucent Technologies*



Digitized by the Internet Archive
in 2023 with funding from
Amateur Radio Digital Communications, Grant 151

<https://archive.org/details/rawcon981998ieee00unse>

PROCEEDINGS



1998 IEEE Radio and Wireless Conference

Sheraton Colorado Springs

Colorado Springs, Colorado, USA

August 9-12, 1998

<http://rawcon.org>



NIST ITS

RAWCON'98 Co-Sponsors

IEEE Microwave Theory and Techniques Society

IEEE Pikes Peak Section

RAWCON'98 Technical Co-Sponsors

IEEE Communications Society

National Institute of Standards and Technology

Institute for Telecommunications Sciences

RAWCON'98 Media Co-Sponsors

Applied Microwave & Wireless

RF Globalnet Website



General Chair:

Dr. Roger Marks, *National Institute of Standards and Technology*

Technical Program Chair:

Dr. Michael S. Heutmaker, *Lucent Technologies*

Proceedings

1998 IEEE Radio and Wireless Conference

IEEE Catalog Number: 98EX194

ISBN: 0-7803-4988-1 (Softbound Edition)
0-7803-4989-X (Microfiche Edition)

Library of Congress: 98-85757

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved, Copyright ©1998 by the Institute of Electrical and Electronics Engineers, Inc.

Additional copies of this publication may be obtained from:

The Institute of Electrical and Electronics Engineers, Inc. (IEEE)
Publication Order Department
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331

Please refer to the Catalog Number or ISBN listed above

Copyright ©1998
The Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved
Printed in the U.S.A.

Steering Committee

GENERAL CHAIR
Roger B. Marks
NIST

TECHNICAL PROGRAM CHAIR
Michael S. Heutmaker
Lucent Technologies

EXHIBITS CHAIR
Michael Fennelly
ATN Microwave, Inc.

REGISTRATION CHAIR
George Jankovic
RF Globalnet Website

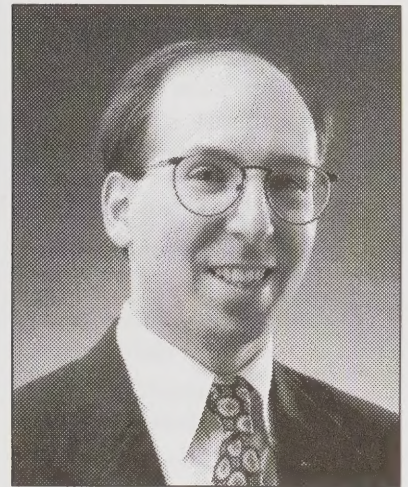
FINANCE
Richard A. Sparks
Anro Engineering
Gerome R. Reeve
NIST

LOCAL ARRANGEMENTS
John G. McIlroy
The MITRE Corporation

PUBLICITY
Thomas T. Y. Wong
Illinois Institute of Technology

PUBLICATIONS
Gary A. Breed
Noble Publishing Corp.

AT-LARGE
Peter W. Staeker
M/A-COM Inc.
John W. Meredith
Hewlett-Packard Co.



Dr. Roger B. Marks
General Chair

This book is the record of the 1998 IEEE Radio and Wireless Conference (RAWCON'98). Although this is the first conference of this name, nearly all of the energy behind it came from volunteers who had participated in the Wireless Communications Conference, held in Colorado in August of 1996 and 1997. The WCC established an international reputation for outstanding presentations by and interactions between many of the world's leading researchers in wireless communications technology. This reputation has apparently carried over to RAWCON, which, as evidenced by the size of this book, has seen great growth compared to WCC'97. Although we had originally planned to continue the format of fifty oral presentations, we were overwhelmed by the quality and quantity of the 149 submissions. Since we were committed to the single-track session format, we found additional space by adding non-overlapping poster sessions. This allowed us to accommodate the approximately ninety superb papers in this Proceedings.

A glance at the papers testifies to the truly international status of RAWCON'98. We received submissions from 27 countries and accepted papers from 22. Nearly half of the papers originate outside the United States.

In a world full of competing events, why has RAWCON'98 proven successful? I think because:

- It provides a strictly peer-reviewed program of new developments with a dedicated international Technical Program Committee, led by an outstanding Technical Program Chair, Dr. Michael Heutmaker of Lucent Technologies.
- It provides an integrated view of wireless technology, from high-level systems down to low-level components, in a congenial and interactive environment. As a result, it fosters customer-supplier as well as peer-to-peer technical relationships.
- It emphasizes critical areas of wireless systems, such as radio frequency aspects, that are sometimes overlooked, and it addresses the important growth areas like broadband wireless.
- It has the support of key organizations in the wireless field, particularly the IEEE Microwave Theory and Techniques Society.

(continued on next page)

- It is organized by an energetic, creative, and technically astute Steering Committee. In conjunction with these Proceedings, I'd like to make particular note of the contributions of the Publications Chair, Gary Breed of Noble Publishing.

Several components of RAWCON that are not recorded here were nevertheless key elements in the conference:

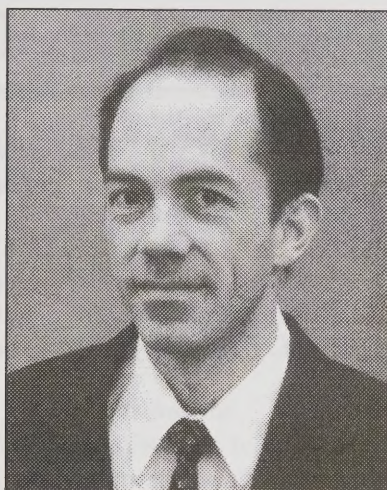
- A Keynote Address on "Technical Challenges to LMDS Implementation" by Barclay Jones of WNP Communications, Inc.
- A Banquet Address by Larry Irving, who, as Assistant Secretary for Communications and Information in the U. S. Department of Commerce, heads the National Telecommunications and Information Administration (NTIA).
- An evening Panel Session entitled "LMDS: Jumpstarting the Industry," moderated by Sanjay Kasturia of Raychem Corp. and organized by Sanjay and me.
- A full-day Sunday workshop on "Modeling and Simulation of Devices and Circuits for Wireless Communication Systems," organized and chaired by John Sevic of Spectrian Corp.
- A half-day Monday tutorial workshop entitled "Overview of Coding and Modulation for Wireless Communications," organized and presented by Prof. Rodger E. Ziemer of the University of Colorado at Colorado Springs.
- An exhibition of products and services for the wireless industry that was organized by Mike Fennelly of ATN Microwave.

I hope that you, either as a conference attendee or simply as a reader of this book, take full advantage of the excellent opportunity that RAWCON offers, both in 1998 and in the future.

Roger B. Marks

National Institute of Standards and Technology

General Chair, RAWCON'98



Dr. Michael S. Heutmaker
Technical Program Chair

On behalf of the RAWCON'98 Technical Program Committee, I welcome you to the 1998 IEEE Radio and Wireless Conference. RAWCON'98 aims to provide a program that covers the breadth of the wireless product development "food chain" in a setting where researchers and technologists can get a better understanding of the RF-related technical issues facing their customers, suppliers, and competitors.

The robust growth in the number of submitted papers (with respect to the 1997 Wireless Communications Conference, from which RAWCON'98 evolved) has precipitated several modifications to the program. The most obvious change is the addition of poster sessions, interspersed with the oral presentation sessions. This scheme was the best way we could think of to accommodate a much larger number of high-quality papers without increasing the overall length of the conference. The result is a solidly-packed three-day program, and we would like to know your reactions to this format.

Much of the growth in the number of submitted papers has come from the "System Performance" technology area. In many cases, system performance issues draw on multidisciplinary aspects of RF technology and so are particularly appropriate for this conference. The Technical Program Committee is trying to anticipate trends in the interests of the RAWCON audience, as well as to maintain strengths that the conference has already established.

I would like to thank each member of the Technical Program Committee for their contributions, especially Kari-Pekka Estola, Peter Petre, Joe Staudinger, and Roger Dalke for their work as sub-committee chairs for the various technology areas.

Michael S. Heutmaker
Lucent Technologies
Technical Program Chair, RAWCON'98

Technical Program Committee

TECHNICAL PROGRAM CHAIR
Dr. Michael S. Heutmaker
Lucent Technologies

COMMITTEE MEMBERS

Robert J. Achatz
ITS

Dr. Masami Akaike
Science Univ. of Tokyo

Dr. Lutfi Albasha
Sony Semiconductor

Dr. Seng-Woon Chen
Polanyi, Inc.

Dr. Roger Dalke
ITS

Dr. Kari-Pekka Estola
Nokia Research Center

Bernhard Geller
Sarnoff Corp.

Dr. David Hill
NIST

Charles Hodges
Flextronics

Dr. Hossein Izapanah
Hughes Research Laboratories

Dr. Luc Martens
University of Gent

Dr. Masahiro Muraguchi
NTT Wireless Systems Laboratory

Dr. Modest Oprysko
IBM T.J. Watson Research Center

Dr. Chong-Il Park
Kyocera America

Dr. Peter Petre
Hughes Research Laboratories

Dr. James Schaffner
Hughes Research Laboratories

Andreas Schmidbauer
Technical Univ. of Munich

Dr. Mohammad Shakouri
Hewlett-Packard Company

Dr. David Smithgall
Lucent Technologies

Joseph Staudinger
Motorola Inc.

Dr. Gene Tkachenko
Alpha Industries

Dr. Vijai Tripathi
Oregon State University

Dr. Mark Wickert
University of Colorado

Special Technical Programs

Sunday, August 9, 1998

8:00 AM - 5:00 PM

**Full-Day Workshop:
Modeling and Simulation of
Devices and Circuits for Wireless
Communication Systems**

Organizer and Chair:
John F. Sevic, *Spectrian Corp.*

Speakers

Jeff Jargon, *NIST*
Steve Maas, *Nonlinear Technologies, Inc.*
James Rautio, *Sonnet Software, Inc.*
David Root, *Hewlett-Packard Co.*
John F. Sevic, *Spectrian Corp.*
Joe Staudinger, *Motorola, Inc.*
Wayne Struble, *M/A-COM, Inc.*
Thomas Brazil, *Univ. College Dublin*

Commercial RF and microwave CAD packages have evolved to the point where predictive simulation is possible. When properly applied, CAD can reduce design cycle-time and provide insight that bench techniques alone cannot. Virtual prototyping and statistical optimization are possible, precluding the necessity of time-consuming pilot runs. Unfortunately, many CAD users are unfamiliar with rigorous modeling and characterization techniques. The result is often user frustration and alienation. Ultimately, many CAD users return to bench methods.

The purpose of this workshop is to introduce basic modeling and characterization techniques for RF and microwave CAD. Experts will discuss how to use simulation to complement the design process for wireless communication circuits. The emphasis will be on learning when to use simulation to an advantage, how to interpret results, and how to establish confidence in the models. Attendees of the workshop will learn to trust simulation in order to make engineering decisions. The wide variety of topics covered includes simulation and modeling of passive lumped and distributed components, frequency-conversion circuits, and power amplifier circuits. Large-signal device modeling and characterization techniques will also be included. The appropriate simulation methods will be identified, and simulations will be compared to measurements.

Monday, August 10, 1998

Opening and Keynote Address

1:00 - 1:40 PM

Co-Chair: Roger B. Marks,
General Chair, RAWCON'98
Co-Chair: Michael S. Heutmaker,
Technical Program Chair, RAWCON'98

Keynote Address
**"Technical Challenges to LMDS
Implementation"**

Barclay Jones,
WNP Communications, Inc.

Monday, August 10, 1998

8:00 AM - Noon

**Half-Day Workshop:
Overview of Coding and
Modulation for Wireless
Communications**

Organizer and presenter:
Prof. Rodger E. Ziemer,
Univ. of Colorado, Colorado Springs

The overall goal of this tutorial is to present modulation and coding techniques important for understanding design trade-offs for wireless communication systems. It is important to understand the advantages and disadvantages that various multiple access, modulation, and coding approaches offer for their implementation. This tutorial will present an overview of the theory necessary for understanding approaches used for second generation systems and those being proposed for third generation systems to be fielded soon after the year 2000.

Monday, August 10, 1998

Reception and Banquet

5:30 - 9:30 PM

Banquet Speaker
Larry Irving
*Assistant Secretary for Communications
and Information, U.S. Department of
Commerce, and Director, National
Telecommunications and Information
Administration (NTIA)*

Tuesday, August 11, 1998

7:00 - 9:00 PM

**Panel Discussion Session:
LMDS: Jumpstarting
the Industry**

Organizers:
Roger Marks, *National Wireless
Electronics Systems Testbed
(N-WEST), NIST*
Sanjay Kasturia, *Raychem Corp.*
(moderator)

Panelists

Marc Leclair,
Baker Creek Communications
Mohammad Shakouri,
Lucent Technologies
Leland Langston,
Raytheon TI Systems
Narisa Chu, *C&W Systems, Ltd.*
Roger Marks, *N-WEST*

With the U.S. LMDS auctions complete, the largest block of spectrum ever licensed will soon be available to provide new services. To fully exploit the enormous potential of this 1.3 GHz of bandwidth, a number of players will have to come together and work cooperatively to address the special challenges of operating in this band. Currently, however, the industry is fragmented at many levels. The licenses are in the hands of dozens of auction winners, many of whom are regional and only a few of which have experience in radio communications. Several major LMDS systems integrators have recently changed ownership and the components required for operation in the 28-31 GHz band remain expensive.

License holders, component suppliers, and potential customers are best served by standardization. However, following recent trends, the U.S. government is leaving technology development and standardization completely to the marketplace. No significant operational LMDS standards exist and, unlike the case of PCS, this industry cannot simply adopt or modify the standards developed from another industry segment. This panel features speakers interested in proactive steps toward a coherent LMDS industry

Table of Contents

MONDAY, AUGUST 10, 1998

Session M1

Broadband Wireless System Concepts

1:40 PM - 3:00 PM

Chair: Modest Oprysko,
IBM T.J. Watson Research Center

- 1 LMDS in the Broadband Local Loop - Alternatives and Issues
Aftab Ahmad, Kieson Kim, KwangJu Institute of Science and Technology, KwangJu, The Republic of Korea
- 5 MM-Wave Wireless Access Technology for the Wideband Wireless Local Loop Applications
Hossein Izadpanah, Daniel Gregoire, Jim Schaffner, H. P. Hsu, HRL Laboratories, Malibu, CA
- 9 Broadband Data Services from a HALO Aircraft
Peter Diamandis, Nicholas Colella, Angel Technologies Corporation, St. Louis, MO
- 15 Transmission Techniques for Increased Power Efficiency in OFDM-Based Wireless Communication Systems
Rui Dinis, Antonio Gusmao, CAPS, Instituto Superior Tecnico, Lisboa Codex, Portugal

Session M2

System Architecture and Networking

3:30 PM - 5:10 PM

Chair: Kari-Pekka Estola, Nokia Research Center

- 19 Cellular Network Design and Smart Antenna Systems
Mark Reudink, Metawave Communications, Redmond, WA
- 23 A New Throughput Analysis of a Novel MAC Protocol for CDMA-Based Wireless LANs
Peter Omiyi, Timothy O'Farrell, The University of Leeds, Leeds, England
- 27 Comparison of FH/CDMA and DS/CDMA for Wireless Survivable Networks
Scott Chuprun, Chad Bergstrom, Motorola Space and Systems Technology Group, Scottsdale, AZ

- 31 Dynamic Router Configuration Management for Wireless Mobile Environments
Sidarth Ambardar, Walter Lazear, The MITRE Corporation, McLean, VA
- 35 A FDM Multiple User Software Configurable Arbitrary Format (MUSCAF) UHF Receiver Prototype
Ian Scott, T/As Minilinx, Perth, Western Australia

Session MA

Wireless/RF Technology (Poster)

5:30 PM - 8:00 PM

Chairs: Mohammad Shakouri, Lucent Technologies;
David Smithgall, Lucent Technologies

System Architecture Posters

- 39 An IF Digitizing Receiver for a Combined GPS/GSM Terminal
K. Boehm, T. Mueller, Daimler-Benz Research Ulm, Ulm, Germany; T. Hentschel, Dresden University of Technology, Germany; F. Oehler, G. Rohmer, Fraunhofer Institute for Integrated Circuits, Erlangen, Germany
- 43 Ultra High Speed Digital Processing for Wireless Systems Using Passive Microwave Logic
Bhaskar Krishnamachari, Simon Lok, Christopher Gracia, Sajan Abraham, The Cooper Union for the Advancement of Science and Technology, New York, NY
- 47 A Wireless Data System Constructed of SAW-Based Receiver/Transmitter and Its Applications to Medical Cares
Kenji Matsumura, Osaka University and KCS Co., Ltd., Osaka, Japan; Gen Fujita, Isao Shirakawa, Osaka University, Osaka, Japan; Hiroshi Inada, National Cardiovascular Center Research Institute, Osaka, Japan
- 51 Analysis and Coloring of Shotgun Cellular Systems
Timothy Brown, University of Colorado, Boulder, CO
- 55 System Capacity, Latency, and Power Consumption in Multihop-routed SS-CDMA Wireless Networks
Matthew Ettus, Sun Microsystems, Mountain View, CA
- 59 A Mechanism for Matching Application Message Generation With Wireless Network Performance
Brian Rivera, Army Research Laboratory, Atlanta, GA; Michael Humphrey, Charlene Todd, Motorola Space and Systems Technology Group, Scottsdale, AZ

System Performance Posters

- 63 N Dimensional Orthogonal QPSK Signaling With Discrete Prolate Spheroidal Sequences
D. M. Gruenbacher, D. R. Hummels, Kansas State University, Manhattan, KS
- 67 System Performance Analysis of Impulse Radio Modulation
Fernando Ramirez-Mireles, Glenayre Technologies, Los Angeles, CA; Robert Scholtz, University of Southern California, Los Angeles, CA
- 71 In-Band Intermodulation Measurement and Simulation of a Local Multipoint Distribution System
Thomas Wong, Illinois Institute of Technology, Chicago, IL
- 75 Noise in Self-timed and Synchronous Implementations of a DSP
W. Lien, P. Day, R. Glibbery, D. Jackson, J. Lin, N. C. Paver, Cogency Technology Incorporated, Ontario, Canada
- 79 Wideband Radio Prototyping Environment
Tommi Makelainen, Timo Eriksson, Jarkko Posti, Risto Suoranta, Nokia Research Center, Helsinki, Finland
- 83 Digital Single Side Band with Pilot Symbol Assisted Modulation in Mobile Radio Channels
Seungwon Kim, Leon Couch, University of Florida, Gainesville, FL

Antenna/Propagation Posters

- 87 Characterization of Bunny-Ear Antennas for Wireless Basestation Applications
Marc Greenberg, Kathleen Virga, University of Arizona, Tucson, AZ
- 91 Effect of Tilted Base-Station Antennas on Cochannel Interference Reduction in Cellular Communications
Ting-Kai Qin, Xin Zhang, Yong Cheng, Nanjing University of Posts & Telecommunications, Nanjing, China
- 95 Electronic Beam Tilting of Base Station Antenna: Rotman Lens Fed
S. G. Kim, S. M. Park, C. W. Ro, J. M. Lee, Gamma Nu Wave Inc., Kyungki-Do, Korea
- 99 An Improved 3D Ray Tracing Method for Indoor Propagation Prediction
W. Lu, K. T. Chan, The Chinese University of Hong Kong, Shatin, Hong Kong

- 103 Miniature Dielectric-loaded Personal Telephone Antennas with Low SAR
Bernhard Rosenberger, Rosenberger HF - Technik GmbH & Co., Tittmoning, Germany
- 109 A Practical Approach to Wideband Antenna Array Characterization
S. London, V. Leonov, P. Koert, L. Susman, Advanced Power Technologies Inc., Washington, DC

Passive Device Posters

- 113 Advanced Chip-in-Board Thinfilm Hybrid Technology with Integrated Polyimide Film Capacitors for Ultra High-Speed and Millimeterwave Applications
Wolf-Dieter Nohr, Gerhard Hanke, Dieter-Jurgen Weber, Deutsche Telekom, Berlin, Germany
- 117 Characterization of Embedded Resistors for High Frequency Wireless Applications
Nanju Na, Kwang Lim Choi, Madhavan Swaminathan, Georgia Institute of Technology, Atlanta, GA

Active Device Posters

- 121 A PHEMT MMIC Broad-Band Power Amplifier for LMDS
Young-Gi Kim, Anyang University, Anyang-City, Kyungki-Do, Korea; Sung-Jae Maeng, Jin-Hee Lee, Chul-Soon Park, Telecommunications Research Institute, Taejon, Korea
- 125 A Two Stage, Monolithic Integrated 200 mW HEMT Amplifier for Wireless ATM
Thomas Bos, Urs Lott, Werner Bachtold, Swiss Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland
- 129 A 85 Mbps Low-Power DQPSK MODEM-LSI for Advanced Wireless Access
Akihiro Yamagishi, NTT System Electronics Laboratories, Kanagawa, Japan; Masahiro Muraguchi, Tsuneo Tsukahara, NTT Wireless Systems Laboratories, Kanagawa, Japan
- 133 A Cost-Effective Approach to a Short-Range, High-Speed Radio Design in the U-NII 5.x GHz Band
Mehmet Soyuer, Herschel Ainspan, Joachim Burghartz, Jean-Olivier Plouchart, Brian Gaucher, Troy Beukema, Frank Canora, Erik Pilmanis, Modest Oprysko, IJB T.J. Watson Research Center, Yorktown Heights, NY
- 137 A High-Efficiency RF Transmitter Using VCO-Derived Synthesis: CALLUM
David Jennings, Joseph McGeehan, University of Bristol, Bristol, United Kingdom

TUESDAY, AUGUST 11, 1998

Session T1

System Modeling and Measurement

8:00 AM - 9:20 AM

Chair: R. E. Ziemer,

University of Colorado at Colorado Springs

- 141 Bit Error Rate Analysis in Wireless Systems Modeling
George Chrisikos, Christopher Clark, The Aerospace Corporation, Los Angeles, CA
- 145 Performance of an Adaptive Rate Modem Using Quasi-Analytic Simulation Techniques
Peter Hofstetter, Mark Wickert, University of Colorado at Colorado Springs, Colorado Springs, CO
- 149 Transmission Capacity of a Broadband Wireless Radio Link
Jose Fernandes, Instituto de Telecomunicacoes - Polo Aveiro, Aveiro, Portugal
- 153 Cellular System Performance Prediction
Ron Rudokas, Littleton, CO

Session T2

System Optimization

9:50 AM - 11:50 AM

Chair: Mark Wickert,

University of Colorado at Colorado Springs

- 157 Toward the 60 gm Wireless Phone
David Smithgall, Lucent Technologies, Princeton, NJ
- 161 Comparison of Antenna Diversity Schemes
B. Gavilanes-Loureiro, I. DeCoster, K.U.Leuven, div.ESAT/TELEMIC, Kard. Mercierlaan, Heverlee, Belgium; E. Van Lil, Fund for Scientific Research (Flanders); F. Perez-Fontan, University of Vigo, Spain
- 165 Spatio-Temporal Equalization for Wireless Communication
Joshua Alspector, John Norgard, John Parish, University of Colorado at Colorado Springs, Colorado Springs, CO
- 169 Interference Resistant Modulation Using Transform Domain Processing
Chad Bergstrom, J. Scott Chuprun, Motorola Space and Systems Technology Group, Scottsdale, AZ

- 173 Smart Channel Assignment Algorithm for SDMA Systems

Flavio Piolini, Anna Rolando, CSELT - Centro Studi e Laboratori Telecomunicazio S.p.A., Torino, Italy

- 177 MAI Cancellation By Optimum Single-User Detection For DS/CDMA Systems Using Random Binary Codes

Naresh Patel, Timothy O'Farrell, The University of Leeds, Leeds, United Kingdom

Session T3

CDMA System Performance

1:00 PM - 2:40 PM

Chair: Seng-Woon Chen, Polanyi, Inc.

- 181 The Application of Low Noise Amplifiers in CDMA Cellular and PCS Systems for Coverage and Capacity Enhancements
Samuel Yang, AirTouch Cellular, Irvine, CA
- 185 IS-95 CDMA Forward Link Optimization Tool
John Payne, Asim Qureshi, AirTouch Cellular, Irvine, CA
- 189 Simulation of Coded MC-DS-CDMA Systems
Thad Welch, United States Naval Academy, Annapolis, MD; Rodger Ziemer, University of Colorado at Colorado Springs, Colorado Springs, CO
- 193 Analysis of Impact on Handset Transmitter Design of the High-Speed Data Requirements in the IS-95-B CDMA Wireless Standard
Giridhar Mandyam, Texas Instruments Inc., Dallas, TX
- 197 Effects of Carrier Tracking in RAKE Reception of Wideband DSSS
Rodger Ziemer, University of Colorado at Colorado Springs, Colorado Springs, CO; L. B. Milstein, University of California at San Diego, La Jolla, CA; B. R. Vojcic, The George Washington University, Washington, DC; J. G. Proakis, Northeastern University, Boston, MA

Session TA

Antenna Design (Poster)

2:40 PM - 3:10 PM

Chair: Robert Achatz,

Institute for Telecommunication Sciences

- 201 Use of Prony's Method for Extracting the Poles and Zeros Yielding a Wideband Window Type Response of Circular Antenna Arrays
Fatma El-Hefnawi, Electronics Research Institute, Cairo, Egypt

- 205 Microstrip Bowtie Patch Antenna for Wireless Indoor Communications
C. H. Ng, S. Uysal, M. S. Leong, National University of Singapore
- 209 Beamforming Applied to an Adaptive Planar Array
Ana Maria Guerreiro, Adriaõ Duarte D. Neto, Fabio Adriano Lisboa, Rio Grande do Norte Federal University, Natal, Brazil
- 213 A Dual-Polarized Microstrip Antenna Array with High Isolation Fed by Coplanar Network
ShiChang Gao, ShunShi Zhong, Shanghai University, Shanghai, P. R. China
- 217 Low-cost, Dual-band, Handset Antennas for LEOS Communications
James McLean, Gentry Crook, Tactical Systems Research, Inc., Austin, TX; Siegfried Mikuteit, Torrey Science Corporation, San Diego, CA

Session T4
Antenna Design and Performance

3:10 PM - 4:30 PM

Chair: Ronald DeLyser, University of Denver

- 221 A Novel Approach for Gain and Bandwidth Enhancement of Patch Antennas
Yongxi Qian, Dan Sievenpiper, Vesna Radisic, Eli Yablonovitch, Tatsuo Itoh, University of California, Los Angeles, CA
- 225 Performance of Lens Antennas in Wireless Indoor Millimeter Wave Applications
Carlos Fernandes, Instituto de Telecomunicacoes - Polo Lisboa, Lisboa Codex, Portugal; Jose Fernandes, Instituto de Telecomunicacoes - Polo Aveiro, Aveiro, Portugal
- 229 Statistical Behavior and Performance of Adaptive Antennas in Multipath Environments
Tien Pham, Mentor Graphics Corporation, San Jose, CA
- 233 On Applications of Self-Phased Array Antennas to Mobile Communications
Sergey Loyka, Vladimir Mordachev, Belorussian State University of Informatics & Radioelectronics, Minsk, Republic of Belarus

Session TB
Signal Processing (Poster)

6:00 PM - 7:00 PM

Chair: Jay Jacobsmeyer, Pericle Communications Co.

- 237 A New Real Time Radio Frequency Direction Finding Algorithm for Gaussian and Non-Gaussian Noise Environments
W. Featherstone, H. J. Strangeways, University of Leeds, Leeds, United Kingdom
- 241 Harmonic Retrieval in Colored Non-Gaussian Noise Using Cumulant and Autocorrelation
Yan Zhang, Shu-Xun Wang, Jilin University of Technology, Changchun, P. R. China
- 245 The use of Arbitrary Sequences for Real-Time Wireless Channel Evaluation
A. Al-Dabbagh, M. Darnell, University of Leeds, Leeds, United Kingdom
- - - Carrier Frequency Tracking for CPM Signal in a Direct Conversion Receiver
Hang Zhang, Nokia Research Center, Helsinki, Finland
- 249 Frequency-Domain Delay Distortion Equalization - The Complex Spectrum Regeneration Method
Masami Akaike, Science University of Tokyo, Tokyo, Japan

WEDNESDAY, AUGUST 12, 1998

Session W1
Propagation Modeling

8:00 AM - 9:20 AM

Chair: Christopher Holloway,
Institute for Telecommunication Sciences

- 253 Frustum Ray Tracing Technique for High Spatial Resolution Channel Characteristic Map
Hajime Suzuki, Ananda Mohan, University of Technology, Sydney, Australia
- 257 VHF Space-to-Earth Radio Link performance in Various Man-Made Noise Environments
Robert Achatz, Roger Dalke, Institute for Telecommunication Sciences, Boulder, CO
- 261 Reception Algorithms for Ray Launching Modeling of Indoor Propagation
Santiago Flores, Luis Mayorgas, Francisco Jimenez, Escuela Universitaria Gandia (Universidad Politecnica de Valencia), Grao Gandia, Spain

- Radio-Channel Propagation Prediction in Urban Micro- and Picocellular Environments
Farshid Aryanfar, J. Rashed-Mohassel, Tehran University, Tehran, Iran; S. Safavi-Naeini, University of Waterloo, Ontario, Canada

Session WA
Passive Device Technology (Poster)

9:20 AM - 9:50 AM

Chair: Charles Hodges, Flextronics, Inc.

- 265 Design Considerations for Extremely High-Q Integrated Inductors and Their Application in CMOS RF Power Amplifier
Tony Yeung, Jack Lau, H. C. Ho, M. C. Poon, The Hong Kong University, Clear Water Bay, Hong Kong
- 269 Microfabricated Single Crystal Silicon Transmission Lines
T. D. Kudrle, H. P. Neves, N. C. MacDonald, Cornell University, Ithaca, NY
- 273 Resonant Frequency of Dielectric Resonator
Juseop Lee, Sang-Min Han, Young-Sik Kim, Korea University, Seoul, Korea
- 277 Design of an Integrated RF Filter for the Direct Digitization Front End of a Dual GPS/GLONASS Software Radio Received
Cornelius Brinegar, Krishna Naishadham, Wright State University, Dayton, OH
- 281 Modeling and Characterization of Wire Bonding and Taped Automatic Bonding. Analytical Formulas and Experimental Validation
N. Hassaine, Y. Shen, P. Ntaka, Harris Farinon Canada, Inc., Quebec, Canada

Session W2
Integrated Passive Devices

9:50 AM - 11:50 AM

Chair: Peter Petre, Hughes Research Laboratories

- 285 Advanced Silicon IC Interconnect Technology: Present Trends and RF Wireless Implications
Ronald Gutmann, Rensselaer Polytechnic Institute, Troy, NY
- 289 Advanced Packaging of Integrated Passive Devices for RF Applications
Elizabeth Logan, Howard Clearfield, James Young, Don Bolton, Intarsia Corporation, Fremont, CA

- 293 Quick Prototyping of Flip Chip Assembly with MEMS
Ronda Irwin, Wenge Zhang, Kevin Harsh, Y. C. Lee, University of Colorado at Boulder, Boulder, CO

- 297 Partially-Interdigitated Combline Filter
Reddy Vangala, Motorola CPG, Albuquerque, NM

- 301 An Improved Lumped-Element Equivalent Circuit for on Silicon Integrated Inductors
Paolo Arcioni, Rinaldo Castello, Luca Perregrini, Enrico Sacchi, University of Pavia, Pavia, Italy; Francesco Svelto, University of Bergamo, Dalmine (BG), Italy

- 305 Spiral Inductor Substrate Loss Modeling in Silicon RFICs
William Kuhn, Naveen Yanduru, Kansas State University, Manhattan, KS

Session W3
Power Amplifiers and Digital Modulation

1:00 PM - 2:40 PM

Chair: Gene Tkachenko, Alpha Industries

- 309 Application of RF LDMOS Power Transistors for 2.2GHz Wideband-CDMA
Alan Wood, Warren Brakensiek, Semiconductor Product Sector, Motorola, Inc., Phoenix, AZ
- 313 Highly Efficient Linear Power Amplifier for 3.5 Volt NADC Cellular Applications
J. Staudinger, G. Norris, R. Sherman, G. Sadowniczak, Motorola, Semiconductor Products Sector, Tempe, AZ
- 317 A New Circuit Topology to Realize High Efficiency, High Linearity, and High Power Microwave Amplifiers
David Upton, Mayflower Communications Company, Inc., Billerica, MA; Peter Maloney, Raytheon Systems Company, Radar & Technology Programs, Sudbury, MA
- 321 GaAs HBT Power Amplifier with Smooth Gain Control Characteristics
Esko Jarvinen, Nokia Research Center, Helsinki, Finland
- 325 Linearization Performance for a Polar Work Function Predistorter
Claudio Rey, Eugene Clark, Spectrian Corporation, Sunnyvale, CA

Session WB
Active Device Technology (Poster)

2:40 PM - 3:10 PM

Chair: Joseph Staudinger, Motorola SPS

- 329 MMIC-Based Class-A/F Power Amplifier for Ultra-Linear-Commercial PCS Multiple-Modulation-Format Base Station Power Amplification
David Helms, Martina Testa, Nelly Vladimirovsky, David Wills, Celwave, Corvallis, OR
- 333 A High Efficiency MMIC Power Amplifier for Phased Array Antenna Applications
Brian Kormanyos, Ronald Kruse, Debra Follensbee, Boeing Information, Space and Defense Systems, Seattle, WA
- 335 Triple Input Single Chip Diversity LNA
Urs Lott, Swiss Federal Institute of Technology (ETH) Zurich, Zurich, Switzerland
- 339 CAD of Microwave Integrated Multistage Active Filters by the Real Frequency Method
P. Jarry, E. Kerherve, M. Hazouard, J. M. Pham, Bordeaux University, Talence, France

Session W4
Active Device Modeling

3:10 PM - 4:50 PM

Chair: Masami Akaike, Science University of Tokyo

- 343 An Examination of Several Large Signal Capacitance Models to Predict GaAs HEMT Linear Power Amplifier Performance
J. Staudinger, M. C. de Baca, R. Vaitkus, Motorola, Semiconductor Products Sector, Tempe, AZ
- 347 Small-Signal and Noise Modeling of Submicrometer Self Aligned Bipolar Transistor
Salvatore Rinaudo, Giuseppe Privitera, Giuseppe Ferla, Agostino Galluzzo, STMicroelectronics, Catania, Italy
- 351 BSIM3 MOSFET Model Accuracy for RF Circuit Simulation
Suet Fong Tin, Ashraf Osman, Kartikeya Mayaram, Washington State University, Pullman, WA
- 355 Table-Based FET Model Assembled From Small-Signal Models
C. J. Wei, Y. A. Tkachenko, D. Bartle, Alpha Industries Inc., Woburn, MA
- 359 Electrical Package Modeling for High Power RF Semiconductor Devices
Tao Liang, Jaime Pla, Mali Mahalingam, Motorola Wireless Infrastructure Systems Division, Phoenix, AZ
-

LMDS in the Broadband Local Loop - Alternatives and Issues¹

Aftab Ahmad *Member IEEE*, Kiseon Kim, *Member IEEE*
{aahmad, kskim}@kjist.ac.kr
Department of Information and Communications
KwangJu Institute of Science and Technology, KwangJu
The Republic of Korea

Abstract

In this paper we look at two important issues in the design of a cellular structure for a local loop offering broadband services. Firstly, we describe how it is possible to use the Local Multipoint Distribution System (LMDS) for distribution services with and without xDSL. Then, we combine the issue of multiple access and frequency planning into a conglomerate of TDMA and CDMA system. We contend that by using TDMA within each cell and CDMA to divide cells not only the limitations of TDMA based systems can be overcome, but also the simplicity of its receiver design is preserved. We also present a cluster based channel allocation mechanism to reduce the interference.

1. Introduction

In an earlier paper [1], we have suggested the use of millimeter wave (MMW) for full services access network (F-SAN) at the physical layer. By definition, the MMW starts at 30 GHz. However, the proposed spectrum is around 28 GHz. Such system has been called as Local Multipoint Distribution System (LMDS) and specifications for one way point to point systems have been proposed by the Digital Audio-visual Council (DAVIC) [2]. The specifications from DAVIC define the physical layer for the intermediate frequency (IF) to give independence of spectrum allocation. As a result of this, each standardization organization is free to allocate whatever part of spectrum it deems feasible. Accordingly, the spectrum allocated differs from country to country. Table I shows some of the allocations.

Country/Area	LMDS Spectrum (GHz)
Korea	25.5 ~ 27.5 (down), 24.25 ~ 24.75(up)
Europe	40.5 ~ 42.5
USA	27.5 ~ 28.35, 31.075 ~ 31.225 (down/up) , 29.1 ~ 29.25 (down)

Table I: Spectrum allocated for LMDS

Our proposal consists of two main parts, namely,
1. How LMDS based system can be used to supplement an existing copper based loop network or, later, an optical fiber access structure.
2. How its propagation properties can be used in its favor to separate community networks from one

another for a frequency plan giving wider area coverage with acceptable outage probability.

In this paper, we go on to the next step and discuss different alternatives and related design issues. Even though our earlier work uses the term Full Services Access Network (F-SAN) instead of broadband local loop, we focus on the broadband aspects of the access network assuming that it is trivial to integrate POTS and N-ISDN with it. In fact, for some time, the largest volume of traffic will consist of only the PSTN like traffic. Internet is in sight to catch up in volume with the PSTN, or even exceed it. Services with movies on demand will jump in as soon as providers of distribution video join the competition.

Given the versatility of the F-SAN, radical design technologies are needed at all levels, the communication and signaling mechanisms (Network level), the communication system design (System level) and even on the infrastructure of subsystems (Device level). However, we will restrict ourselves to the network level, that too on two issues, the physical configuration and multiple access. We will make a new proposal to consider the effect of frequency planning on multiple QoS services.

2. Network Architecture

In any broadband band access network, many issues are to be dealt with regardless of the nature of the medium. Certain features peculiar to the wireless local loop are widely being denoted. Among these, we will choose the following three for this paper:

1. Physical Layout
2. Multiple Access Mechanism

¹ This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks Research Center at Kwangju Institute of Science and Technology

3. Frequency Planning

In the following, we will touch upon these one by one.

2.1 Physical Layout of the Access Network

The main incentive for the FSAN² group (previously called the GX initiative) for using copper is that the current loop networks are almost mostly copper based and the members of FSAN would like to put these wired networks into reuse. Accordingly, the initial deployment of F-SAN may have optical fiber terminating in an Optical Network Unit (ONU) in a neighbourhood node. The ONU, in this way, also becomes the hub of the customer premises network and is supposed to handle many functions in both directions of communication. These functions may include:

1. Bandwidth management in the reverse link without causing any traffic reshaping
 2. Termination and conversion of the electrical signals to optical on the reverse channels and vice versa on the forward channel.
 3. Distribution of video channels in the forward direction
 4. Other supplementary functions, such as ranging etc.
- Owing to the requirement of LoS transmission, the best configuration for LMDS is the same, that is, an LMDS link from the head-end node (HEN) to a neighbouring node (NN). From the NN, a copper distribution network can use VDSL technology to the Customer Premises Network (CPN). However, for a new provider of VoD services other options are also open, such as, an LMDS point to point link to a NN followed by a point-to-multipoint LMDS link to replace VDSL loops. The difference of the two options are obvious as both have pros and cons.

2.2 Multiple Access Mechanisms

All multiple access mechanisms can be grouped into one of the two Resource Division Multiple Access (RDMA) categories. These are Separate Access and Transmission (SAT) and Combined Access and Transmission (CAT). CAT has been popular with networks providing only data services while SAT is almost the only solution for QoS guaranteed networks which can combine isochronous traffic sources with delay insensitive data traffic. As a result of this, all current proposals have request channels for resource reservation. For the request channel pure or slotted Aloha is again the only choice as the number of subscribers are usually unspecified in wireless

networks. Therefore, the multiple access mechanism may consist of two phase, namely,

1. the request and
2. reservation

In the request phase a station is sure of resources only probabilistically while in reservation mode, the bandwidth is available for transmission of information.

The alternatives for resource division are time division, code division, or a combination. While TDMA is the simplest and perhaps more mature technology, it has the disadvantage of wasting bandwidth in providing frequency reuse guard. CDMA, on the other hand does not require a strict frequency planning, but may require stringent power control mechanism resulting in a relatively complex transceiver. The joint or hybrid schemes, however robust they may appear, only increase the complexity. It must be emphasized that the transceiver complexity is one of the factors deciding the success of the F-SAN. The complexity, therefore, must be considered, at least at the beginning. The CDMA technology cannot be ignored altogether with the rich options it has off its sleeve. So, our resolution is that the RDMA should be in the form of TDMA. We will use CDMA in a different way, as shown in the next section.

2.3 Frequency Planning

One of the main advantages that accompany the CDMA system is the frequency planning for reuse. A simple analysis of a TDMA system will easily show that the cell structure depends upon QoS of the applications. Then, with a system providing multiple QoS, frequency distribution among neighbouring cells becomes a tricky issue. CDMA avoids this co-channel interference problem by clubbing together the interference from the same and neighbouring cells and combating it by using autocorrelation for signal detection. Given the bandwidth needs of the wireless F-SAN, and the need of multiple QoS services a CDMA system seems to be more appropriate. This aspect can be incorporated into a WLL by dividing cells based on a unique PN code for each cell. This is desirable because of many reasons. These are:

1. The whole bandwidth can be used for adjacent cells.
2. Within a cell, TDMA will make the transceiver design simpler.
3. The system designer does not need a large number of codes for this purpose. As many PN codes are necessary as the frequency divisions in a TDMA system. Such a small number of PN codes can be found very easily and be used semi-permanently for each cell.
4. If, at a later stage, mobility is added, the soft hand-off is an added benefit, which is required for certain applications.

² We differentiate between the FSAN group and Full Services Access Network (F-SAN) by using the hyphen in the later

Such a frequency planning entails the advantages of both the TDMA and CDMA avoiding the co-channel interference of both. A centralized channel allocation algorithm can reduce the interference further as shown in the following.

2.3.1 Forced Interference Reduction

In this method, the channel allocation is not performed on per cell basis, but on per cluster basis. A cluster is a group of cells which can cause co-channel interference to a given cell. An example of cluster is the seven cells forming one frequency block in TDMA system. Refer to Figure 1 below.



Figure 1: Typical cell cluster in TDMA

In TDMA systems, cell number 1 to 7 use different parts of the cellular spectrum while cell numbers i, j, k and m constitute the interfering channels. In CDMA, calls in each cell cause intra-cell interference to all other calls in the same cell, while the calls in all other cells contribute to the intercell interference. In contrast to both these interference types, in our scheme,

1. There is no intracell interference just as in TDMA
2. There is less intercell interference than CDMA
3. Interference in pure CDMA system can only be controlled by power control while in our scheme, there is one other measure possible to control the intercell interference. This measure, which we call Forced Interference Control (FIR) is possible as described below.

We treat the block of cells with significant interference as one cluster. Let c be the cluster size. If there are m channels per cell, then the total number of channels in a cluster is $c \times m = cm$. The nature of interference is such that only calls using the same frequency channel will cause interference (as in TDMA). The channels are allocated centrally by a cluster controller. This cluster controller stores a $m \times c$ matrix of channel numbers in columns and cell number in rows, such as the one shown in Figure 2.

The FIR algorithm will, for a given column, pick up a row with the least number of allocated channels. A cell allocation example is shown in Figure 3. In this diagram, the channels in use are shown as shaded.

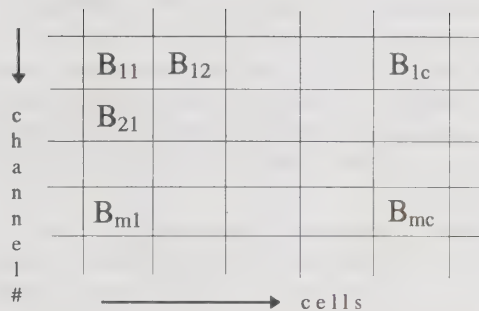


Figure 2: Channel Allocation Matrix (B_{ij} = channel # i of cell # j)

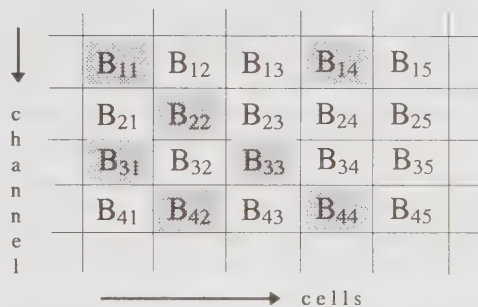


Figure 3a: Before a call arrival in cell 5

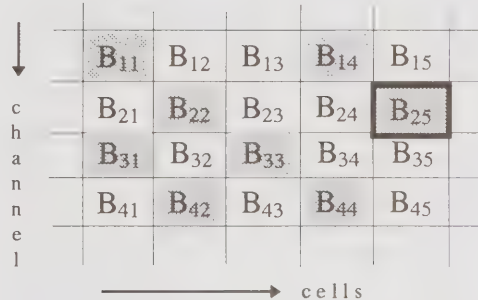


Figure 3b: Channel allocation in cell 5

In this Figure, there are 5 cells and 4 channels per cell. Cell number 1, 2 and 4 have 2 calls in progress while cell number 3 has one call and number 5 has no call in progress. Let's say a new call arrives in cell number 5, then the cluster controller will search column number 5, to make sure that there is a channel available. After that, it will compare the rows of all the entries corresponding to the available channel numbers. Then it will allocate that channel whose rows has the minimum entries.

The implementation complexity of this method can easily be determined from the fact that as the calls come and go, the central processor maintains a state vector for each row. It is required to maintain c such vectors. On arrival of a new call, search is done on the corresponding column. Then the row with minimum entries is chosen.

Interference Reduction with FIR: Usually, communication system design is carried out for worst case conditions. However, in actual practice worst case conditions occur with rarity. The occurrence of such condition is regulated by the standardization agencies in terms of blocking probability, which is of the order of a few percent. For major fraction of system operation, an average condition of traffic exists. The improvement by FIR can be estimated by noting that the first co-channel interference will occur when there are $m+1$ calls in a cluster. The interference remains the same for each new call until there are $2m$ calls, after which there will be 2 interfering calls, and so on.

Thus if, at an arbitrary time, there are N calls in the cluster, each causing an interference of S units. Then we can form the expression:

$$N = km + r \quad \dots 1$$

where $k = 0, 1, \dots, c$ and $r = 0, 1, \dots, m-1$

In FIR system, the interference is kS or less, while in CDMA, the interference is $(N-1)S$.

In practice S is a stochastic quantity and a more rigorous analysis will be required to see the exact benefit. This example, however, shows the source of benefit, the selective allocation.

It is a well-known fact that interference is not the only measure for making a system design decision. Other factors, such as, user capacity have to be evaluated.

FIR in TDMA Systems: FIR can be applied on the TDMA system using the TDMA based frequency planning. However, such a system is not suitable for multimedia services requiring different QoS [3].

FIR in CDMA Systems: In pure CDMA systems FIR is not possible as every channel uses up the whole bandwidth.

We will end this section by saying that a layered TDMA/CDMA access and frequency planning mechanism would best suit the needs of an F-SAN based on LMDS. There are other proposals too. One of particular interest is proposed in [4].

3. Concluding Remarks

In this paper, we have discussed the physical configuration, multiple access and frequency planning issues for a broadband local loop using LMDS on the physical layer. In view of the limiting propagation properties of the MMW, a LoS transmission to a neighbourhood node is required. For simple receiver structure, a TDMA based system is arguably the best.

However, it may be said that TDMA technology is not suited to the multimedia services owing to the cell structure based on QoS. This problem can be overcome by dividing cells based on a unique PN code. This layered architecture can use some additional processing for interference reduction. We have shown that by using a proper mechanism of channel allocation for a cluster of cells, extra reduction in the interference is possible. This reduction will result in extending the battery life, a much desirable feature in wireless systems.

Regarding the issues of the communication system design and the infrastructure of subsystems, it is noteworthy that cost-efficiency issues are extensively under studying [5]. Communication systems for local loop subscribers need to be economic, and spectral efficiency of the broadband system is, yet, an important issue, apparently it enjoys enough bandwidth resources though. Also, considering the integration of subsystems on chip, specially handling 155Mbps-class services, cost-efficiency of the recursive-type algorithms are to be replaced by simple feed-forward ones. Compared to the complexity due to the variations of the bandwidth efficient modulation system, the algorithmic simplifications contribute more significantly to the cost-efficiency issues. Likewise, related to the economic deployment of the physical millimeterwave infrastructure, the emerging device technology is under developing to afford enough margin and link availability against foliage shadowing and expected precipitation rate, at a economic price to the customers.

References

1. Ahmad, A., Kim, K., Song, J., "Network Architecture for HFR Access Network with Millimeter Wave", *LANMAN'98*, Banff, Canada, May, 1998
2. DAVIC 1.1 Specification Part 08 revision 3.0, "Lower Layer Protocols And Physical Interfaces", Digital Audio-Visual Council
3. Lee, W.C., *Mobile Communications Design Fundamentals*, John Wiley & Sons, 1993
4. Cornaglia, B. and D'Aria, G., "Radio Systems Architectures for Wireless CATV", *ISSLS 96* pp 128-133
5. K.Kim et al, An Introduction to Broadband Digital CATV System Technology, Workshop Note, Seoul, Oct., 1997, in Korean

MM-Wave Wireless Access Technology For The Wideband Wireless Local Loop Applications

Hossein Izadpanah*, Daniel Gregoire, Jim Schaffner, and H. P. Hsu
HRL Laboratories, Malibu, California

* Tel: (310) 317-5563, Fax: (310) 317-5485, Email: hizad@hrl.com

Abstract

In this paper, we describe and demonstrate a Wideband Wireless Local Loop (W-WLL) testbed concept based on LMDS millimeter wireless technology but with a modified access and backbone architecture. The testbed is intended as a demonstration platform for broadband wireless services particularly for high speed internet and shared multimedia applications. The investigation is focused on the radio link design, network architecture, system integration, and compatible interface to the existing ATM fiber and satellite core networks.

Introduction

Worldwide, the demand for greater bandwidth is exploding, creating a bottleneck in the user access and at the last mile of the network. Recent spectrum auction and the opening of spectrum (27-31GHz band) for Local to Multipoint Distribution Services (LMDS) and the availability of the 5 GHz U-NII band, are accelerating wireless broadband network implementations. Traditionally LMDS and MMDS (Multi-channel Multipoint Distribution Services) have been used to deliver video entertainment programs and, for institutions, to deliver video for distance learning activities in the analog format. However, the technology is undergoing fast transition to adopt digital video and video compression for spectral efficiency and increased system capacity. These digital LMDS system and the large usable bandwidth at mm-wave frequencies can now offer data rates at up to several Gbps for the W-WLL. The system is capable of supporting full-duplex, high data rate applications such as high-speed Internet, interactive video, video conferencing, and simultaneously including hundreds of traditional broadcast and digital television channels. The LMDS service operates in a cellular fashion, transmitting information between a central hub and the subscriber fixed units installed at businesses and homes using highly directional antennas. Interconnection of hub sites and hubs to the central offices also make use of the mm-wave technology and/or fiber backbone networks.

The W-WLL benefits from the LMDS technology including cost effectiveness, fast

flexible radio deployment, and low maintenance. However, the conventional LMDS system concept and architecture, originally designed for one-way distribution, are not optimized for a two-way broadband applications and requires modifications and parameter optimizations in order to guarantee a higher reliability and increased link availability.

In this paper, we propose and demonstrate a short-range line-of-sight LMDS-like architecture for the W-WLL system which possesses many technological and operational advantages. These include ease of installation and alignment, low radiation power, and effectively, the link being free from major multipath, obstructions (trees, buildings, and moving objects), and adjacent cell interference. The proposed system architecture is highly suitable to be deployed in a highly developed, densely populated urban environment where large capacity broadband services are in great demands but lacks a broadband access infrastructure.

System Description

The new architecture for W-WLL access and backbone connections is rapidly deployable at relative lower cost, in communities lacking suitable infrastructure or where separate broadband infrastructures do not exist to the end users for integrated telephone, internet and video services. The proposed mm-wave access network is a cellular based point-to-point and/or point-to-multipoint distribution system which resembling the traditional LMDS system but, with the following system modifications:

- ⇒ Convert the conventional LMDS standard size cell of 2-5 miles in diameter into smaller micro/pico size cells. Examples of service scenarios are customers concentrated in small areas such as college campuses, business parks, multi-story/high-rise buildings, or planned housing complexes and development in small communities.
- ⇒ Establish the “Hub” access by direct line of sight “illumination” say the multi-story building faces either rooftop and/or sidewall mounted shower type antennas as shown in Fig. 1. The campus and small community access can be provided by projecting the signal from the antennas mounted on the street lamp posts. See Fig.1
- ⇒ Use the spectrum availability to implement multi-user, bi-directional, broadband service links and, if desirable, integrate with the traditional broadcast TV channels for combined multi-services operation. The possibility of these combined services through a single infrastructure also leads to reduced cost for the service providers
- ⇒ It is further proposed to use the existing embedded fiber to the neighborhood (FTTC or HFC now available most service areas) for the broadband backbone network integration.
- ⇒ Relocate the hub radio control and switching facilities to the central office or switching centers for an integrated operation. The consolidation will benefit for lower infrastructure complexity and cost resulting in a more reliable and centralized data base and operations.

New System Architecture Advantages

Compared to the traditional LMDS system [1], the proposed direct LOS distribution topology possesses many technological and operational advantages. These include:

- ⇒ Increased coverage and user penetration percentage in each individual cell due to densely positioned users in the service area. This relaxes the tedious effort of cell frequency and polarization reuse planning, overlapping cells design, and cell boundary coverage planning.
- ⇒ The required Hub and customer transmitting power (at mm-wave!) are immediately scaled

down (15-dB minimum), due to the relatively short cell diameter. The result is a low power low cost system solution and less complex MMIC hardware design.

- ⇒ A major reduction in the system interference (adjacent channel and adjacent cell) constraints and limitations due to the higher power amplifiers nonlinearities
- ⇒ Possible reduction in the required radio channel spacing leading to an increased system capacity
- ⇒ The near short-range LOS propagation path becomes free from “major” multi-path interference, inter-cell interference, and obstructions (buildings, moving objects, trees, and foliage). Consequently, the propagation path loss approaches to that of square law leading to power efficient system
- ⇒ Improvement in the system gain margin (7-10 dB) and link availability comes from the short line of sight distance which removes the signal reception limitation due to excessive rain attenuation and system down time experienced in higher power LMDS systems.

Preliminary Testbed Results

At HRL, we are implementing of an experimental broadband WLL access links and testbed based on the above proposed concept. A modular multi-band RF architecture including the unlicensed ISM band at 5 GHz and millimeter wave frequencies of 28 and 41 GHz is being implemented using commercial off-the-shelf (COTS) components. The key issue, in the topology described here, is that the Hub transmitter has a low power (< 10 dBm at 28 GHz) with the subscriber return path transmitting power of ~ -4 dBm practical for mass deployments. The common IF circuitry provides an interface between the high-speed baseband modem and the microwave/mm-wave upconverter modules as shown in Figure 2-a. The high-speed (>5 Mbps and up to 30 Mbps) commercially available burst modem subsystem uses a QPSK or M-QAM modulation which conforms with the LMDS and W-WLL modulation format [2]. The subscriber units are fixed installations utilizing COTS developed for mm-wave LMDS with low cost RF and MMICs for commonality and cost reduction. Typical Customer Premises Equipment (CPE) functionalities for a single user and the 28 GHz down converter are shown in Figure 2-b.

In an initial small-cell testbed experiment, we demonstrated that a DBS signal (of 80 digital, 16 transponders, QPSK broadcast TV channels) could be successfully distributed from a central hub to several line-of-sight subscriber units located at various distances from the central hub. The DBS information was upconverted from its 950-1450 MHz baseband for transmission at a mm-band centered around 28 GHz. Measurements confirmed that the signal strength to the subscribers fell off according to the $1/r^2$ law. Further measurements along paths distorted by multipath or partial obstructions such as tree limbs deviated from the $1/r^2$ in a complex manner due to the phase interference. The DBS signal received at the subscriber units was successfully converted into NTSC for television display. As the testbed continues to be developed, it will allow for transmission back from the subscriber units to the central office. The two-way communication for the mm-wave access will rely on a Frequency Division Duplex (FDD) scheme.

We then demonstrated that the central hub could successfully transmit multimedia content in the form of a combination of the DBS signal stacked with broadband (~30 mbps) bi-directional channels. These were combined at the baseband before being upconverted to the mm-wave access band. In this preliminary demonstration, the IF waveforms are represented by unmodulated carrier waves at 1800, 1825, 1850, and 1900 MHz as seen in Fig. 3. Eventually, each of those channels will be implemented using Frequency Division Multiple Access (FDMA) techniques to transmit user-specific voice/data information to its targeted subscriber unit. Figure 3 shows the multimedia waveform driving the central hub 28 GHz upconverter (a), and at the output of the receiver downconverted at the subscriber unit 30 meters away (b). The expanded waveforms for the Hub transmitter and 4-tones are shown in Fig. 3 (c) and (d), respectively.

If the subscriber units do not have the advantage of a line-of-sight location with respect to the central hub, or if nearby obstructions creates substantial multipath interference, then the testbed performance is compromised. Figure 4 illustrates the effect of an indoor multi-path propagation when the hub and subscriber units are separated by 26 m at either end of a narrow (2 m wide by 3 m high) hallway. The proximity of the hallway walls to propagation path leads to a severe multipath condition, that is exhibited as a extreme variation in the signal level across the frequency bandwidth. Traces 4b and 4c show the effects of placing a person in the hallway at distances of 8 and 16 meters from the transmitter, respectively. Compared to Fig. 4a, the presence of this obstruction and its particular location substantially changes the multipath components at the receiver and it is seen in the spectral variation in the received signals.

Conclusions

We have introduced and demonstrated a short-range directly projected fixed wireless point-to-multipoint LMDS system architecture for W-WLL applications capable of carrying enhanced voice, data, and video services. The initial results obtained from a 28 GHz indoor/outdoor testbed is proving the practicality of the system for a low cost reliable broadband loop applications with a differential advantages over the conventional system.

1. Gray, "A broadband wireless access system at 28 GHz", Paper M1.1, Wireless Comms Conference, Boulder CO, August 1997.
2. D. Raychaudhuri et al., "WATMnet: A prototype Wireless ATM System for Multimedia Personal Communication", IEEE JSAC, vol 15, No 1, pp 83-95, Jan. 1997.

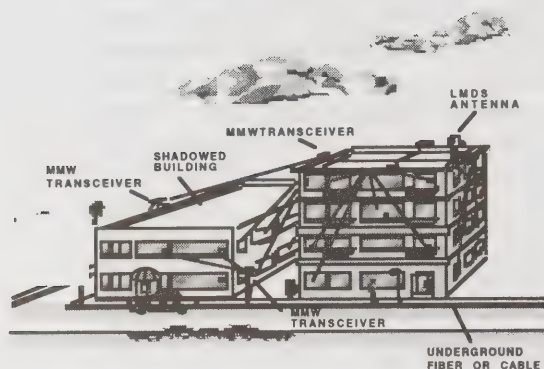


Figure 1- Broadband Fiber Network Fed mm-Wave Distribution To Office Building.

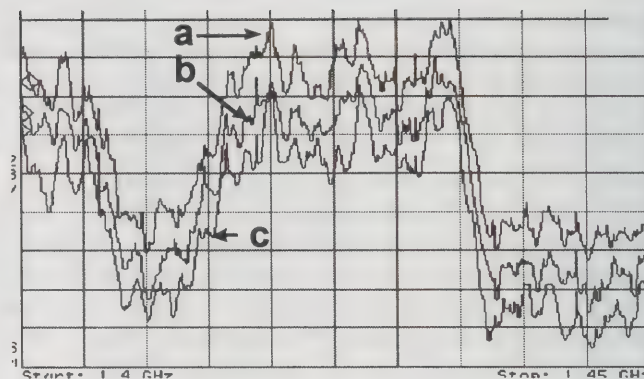


Figure 4 - Recorded 26 m Indoor Multipath

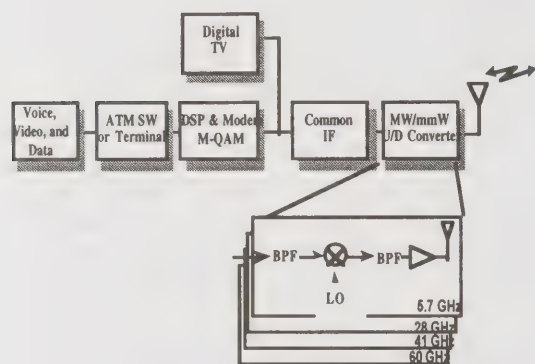


Figure 2 - A Modular Multi-Band RF Architecture For W-WLL Testbed

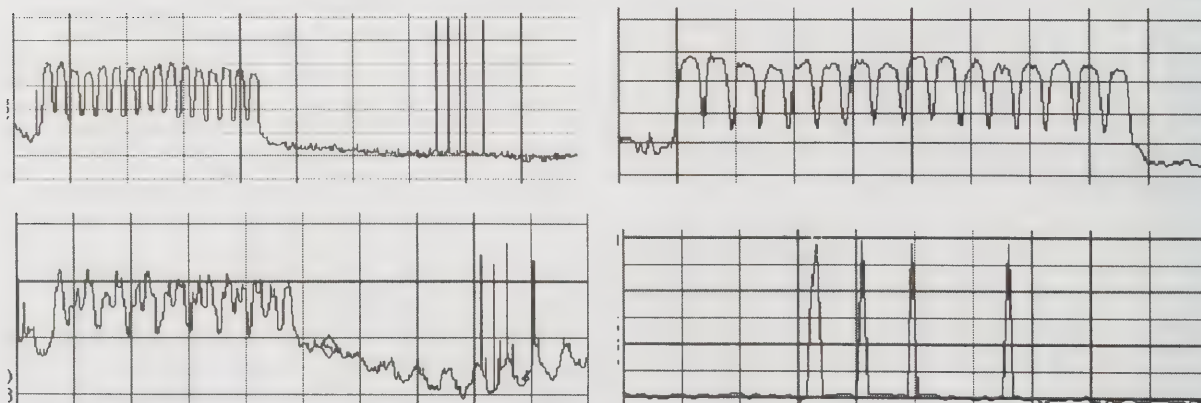
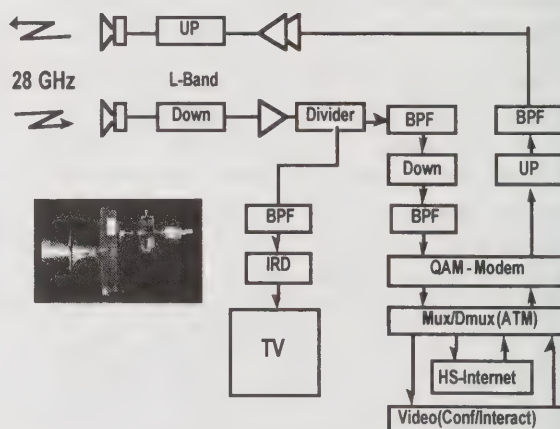


Figure 3 - Mixed signal 950 MHz Wide MM-Wave Transmit/Receive Modulation Spectrum

Broadband Data Services from a HALO Aircraft

Peter H. Diamandis and Nicholas J. Colella

Angel Technologies Corporation - 314-918-1700/phone 314-918-1710/fax - www.angelcorp.com

ABSTRACT

Scaled Composites Incorporated, a partner of Angel Technologies Corporation (www.angelcorp.com), has recently built and will soon flight test a High Altitude Long Operation (HALO) airplane specially engineered for providing a wireless communications networks. HALO aircraft has a fixed-wing airframe with twin turbofan propulsion. It will be FAA certified for piloted operation. Its high altitude flight regime was pioneered in the 1950's and its components will benefit from decades of aerospace industry experience and innovation.

The HALO Network will serve tens of thousands of subscribers within a super-metropolitan region, by offering ubiquitous access throughout the network's signal "footprint". The HALO aircraft will carry the "hub" of a wireless network having a star topology. The initial HALO Network is expected to provide a raw bit capacity exceeding 16 Gbps, which by utilizing packet-switching could, for example, serve 50,000 to 100,000 subscribers requiring links with DS1-equivalent peak data rates in both directions. Three HALO aircraft will fly in shifts to provide continuous service, 24 hour per day by 7 days per week, with an overall system reliability of 99.9% or greater. The HALO airplane will fly above commercial airline traffic and adverse weather at altitudes higher than 51,000 and will provide a communications service footprint or "Cone of Commerce" of approximately 120 kilometers in diameter. Any subscriber within that region will be able to access the HALO Network's ubiquitous multi-gigabit per second "bit cloud" upon demand.

Keywords: HALO Aircraft, HALO Network, Cone of Commerce™, broadband wireless services, metropolitan area network, bit cloud.

1.

1. INTRODUCTION

An electronic "information fabric", a weave of space-based, atmospheric, and terrestrial data communications infrastructure, is emerging that will link digital information devices anywhere on the Earth. Packet-switched data networks will meld with or replace connection-oriented telephony networks. An era of inexpensive, ubiquitous bandwidth has begun.

The convergence of market demand, regulatory reform, innovative technologies and manufacturing capabilities affecting aviation, millimeter wave wireless, and multi-media communications industries enables Angel to pursue new wireless broadband communications services. The HALO Network will offer ubiquitous access to any subscriber within a "super metropolitan area" from an aircraft operating at high altitude. The aircraft will serve as the hub of the HALO Network serving tens of thousands of customers. Each customer will be able to communicate at multi-megabit per second data rates through a simple-to-install subscriber unit. The HALO Network will be steadily evolved at a rapid pace defined by the emergence of advanced technology worldwide. The HALO Network will be a wireless communications network solution. It is an evolving template to be deployed globally on a city-by-city basis.

2. THE HALO NETWORK

Subscribers in homes, businesses, schools and health-care facilities will be able to communicate with each other and with producers/distributors of information, special content, and entertainment services. The subscriber's premise equipment will be standards-based and will require a level of understanding by the subscriber no higher than an ordinary modem for today's personal computers. A small dome mounted on the outside of the premise will house a high-gain antenna with a small aperture that will automatically track the HALO aircraft circling miles away at high altitude.



Figure 1: Premise Equipment for The HALO™ Network

At the apex of a wireless Cone of Commerce™, the payload of the HALO aircraft will serve as the hub of a star topology network for routing data packets between any two subscribers possessing premise equipment within the service coverage area. A single hop with only two links is required; each link connecting the payload to a subscriber. The links are wireless, broadband and line of sight.

Information created outside the service area is delivered to the subscriber's consumer premise equipment ("CPE") through business premise equipment ("BPE") operated by Internet Service Providers ("ISPs") or content providers within that region, and through the HALO Gateway ("HG") directly connected to distant metropolitan areas. The HG is a portal serving the entire network. It avails system-wide access to content providers and it allows any subscriber to extend their communications beyond the HALO Network service area by connecting them to dedicated long-distance lines such as inter-metro optical fiber.

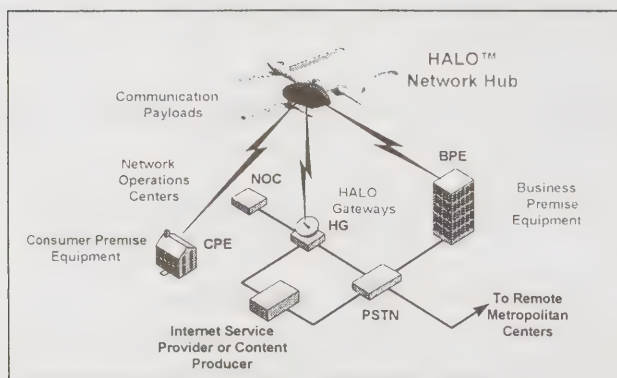


Figure #2: The HALO™ Network

The CPE, BPE and HG have functions in common. They use a high-gain antenna that automatically tracks the HALO aircraft; extract modulated signals conveyed through the air by millimeter waves; convert the extracted signals to digital data; provide standards-based data communications interfaces; and route the digital data to information appliances, personal computers, workstations and servers connected to the premise equipment. Thus, some of the technologies and components, both hardware and software, will be common to the designs of these three basic network elements. All of these elements are being demonstrated in related forms by terrestrial 38 GHz and 28 GHz (LMDS) vendors.

As with all wireless millimeter wave links, rainfall will reduce the strength of the signal received by a given subscriber. For customers at the perimeter of the signal footprint who own CPEs, Angel plans to ensure maximum data rates more than 99.7% of the time, provide reduced data rates above an acceptable minimum more than 99.9% of the time, and to limit outages to small areas (due to the interception of the signal

path by very dense rain columns) less than 0.1% of the time. Customers at shorter slant ranges will have higher availability. Angel plans to locate the HG close to the HALO orbit center to reduce the slant range from its high-gain antenna to the aircraft and hence the length of the signal path through heavy rainfall.

Angel assumes the minimum look angle ("MLA") will be higher than 20 degrees typically, perhaps as low as 10 degrees for regions with low rainfall attenuation. The MLA corresponds to subscribers at the perimeter of the service footprint and is defined to be the elevation angle above the local horizon to the furthest point on the orbit as measured from the antenna of the premise equipment. Angel chose such a high look angle to ensure that the antenna of each subscriber's premise equipment will very likely have access to a solid angle swept by the circling HALO aircraft free of dense objects, and to ensure high availability of service to all subscribers even during heavy rainfall.

A frequency reuse approach similar to traditional cellular mobile networks can be employed in the HALO™ Network. The available spectrum would be divided into frequency sub-bands that are projected from the communications pod as separate "beams". Details on this approach are also described in prior papers, "Broadband Wireless Services from a HALO™ Aircraft" and "The Cone of Commerce".¹

Suppose the frequency bands providing the up and down links to the HALO™ aircraft are divided into five (5) equal sub-bands {A, B, C, D, and E}. A subset of four (4) sub-bands {A, B, C, and D} can then be arranged to form a square tile which, in turn, is tessellated to cover the service area. Each sub-band within a tile, whether labeled A, B, C, or D, will serve many subscribers who share the capacity available in the beam spot corresponding to that sub-band.

On the other hand, sub-band {E} is "orthogonal" in frequency to the other four (i.e., its spectrum is disjoint of the other four sub-bands). It offers frequencies not used in any tiles. Consequently, it can be located anywhere within the signal area to serve premiere customers on a dedicated basis. (This assertion requires the separation between neighboring {E} beam spots be not less than the minimum distance linking the centers of neighboring tiles within the service area, assuming all sub-bands involve similar size antenna apertures on the HALO™ aircraft.)

Each frequency tile repeated on the ground corresponds to a single airborne antenna "cell" in the pod suspended beneath the HALO™ aircraft. The antenna cells are arranged somewhat like the photo-receptors forming the eye of a common household fly to cover the service area. Each cell creates many beams, i.e., multiples of the four sub-bands {A,B,C,D}, through a common refractive optical train specifically optimized for millimeter wave throughput.

Whereas, the {E} beam spots are created by conventional, high-gain antennas that are pointed and stabilized by electro-mechanical or electronic steering means.

3. THE HALO AIRCRAFT

The HALO aircraft fuselage contains the Airborne Switching Node ("ASN"), the primary coolant loop, and power conditioning. Packing switching and network management functions are performed by the ASN. The communications pod suspended beneath the aircraft fuselage contains the millimeter wave antenna array with amplifiers and transceivers. It converts millimeter waves to and from digital signals and is composed of an array of antennas that beam signals to subscribers and to the HG. Power and coolant flow between the platform and the payload through a pylon mount which, in turn, can maintain the payload pod level relative to the ground about the aircraft roll axis if required. A standard optical interface conveys digital communications data across the pylon interface that connects the ASN to the Mux/Demux circuitry in the Pod, which, in turn, impresses the modulated signal upon or extracts it from the millimeter wave carrier.

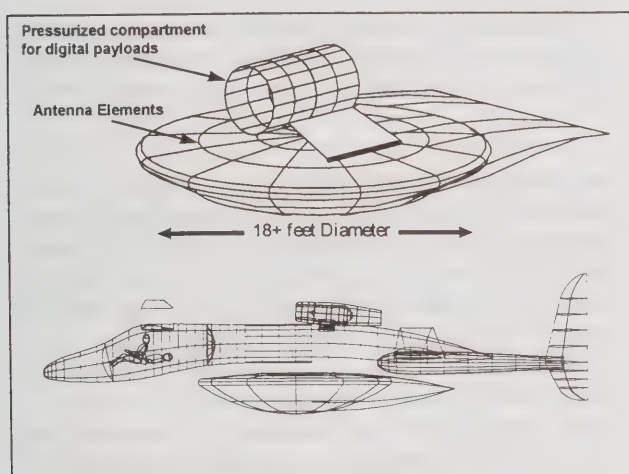


Figure #3: HALO™ Aircraft with Its Communications Pod

From a payload perspective, the HALO™ aircraft is a "flying antenna". It is being prepared for flight testing with Burt Rutan and his team at Scaled Composites during the Summer 1998.

The aircraft configuration is fixed wing with carbon composite materials. Propulsion will be provided by FAA-certified twin fan jets. The platform will be operated by a pilot and co-pilot. The crew cabin will be pressurized to provide a comfortable "shirt sleeve" work environment. Station keeping will be performed by an auto-pilot utilizing coordinates provided by the Global Position Satellite constellation. The two pilots will alternate the duty of flying the airplane. The pilot who is relieved of flight responsibility can be made available to perform routine status checks of the communications payload, including the operations of the components of the

ASN in the fuselage section, as well as those in the suspended Pod. Smart diagnostics data regarding health, status, and operation of the communications equipment will be graphically displayed to the pilot to enable early detecting and forecasting of problems. Similar but wider data streams will be sent to the local operations center on the ground through the Gateway to enable technicians to monitor the status, health, and performance of the airborne communications node.

The payload bearing attributes of the HALO aircraft include:

- 2300 lb. for the airborne elements of the HALO Network distributed as 1800 lb. for the suspended communications Pod, and 500 lb. for the fuselage section with the pylon and the ASN;
- 18 ft diameter communications Pod housing the antennas and their interfaces;
- 40 KVA power for the ASN and Pod elements with matched thermal management; and
- If required, bank angle compensation of the Pod.

4. CONCEPT OF OPERATIONS

The HALO aircraft will execute a circular orbit of roughly 5 to 8 nautical miles in diameter at an altitude above 51,000 feet. From that vantage, broadband services can be offered to users over an area of several thousand square miles or larger encompassing a typical major city and its surrounding towns and communities.

Piloted Aircraft. In order to streamline the aircraft development and FAA certification processes, the HALO aircraft will be manned with two pilots. Over time, Angel may decide to transition to single pilot operation, to be followed by unmanned operations, if the regulatory climate becomes more favorable.

Number of Aircraft per Site. Angel will maintain a continual HALO aircraft presence on-station above each market served. Angel has planned for 8 hour mission times on station, although the airplane is capable of much longer endurance. Three aircraft will be assigned to each isolated site: one on-station, one ascending or descending or being prepared for flight at the local or regional airport facility, and one spare.

Rapid Ground Turnaround Time. An emphasis on "line replaceable unit" maintenance will allow Angel's dedicated personnel to perform regular and unscheduled maintenance actions between missions.

Weather Diversions. Angel will pursue a certificate under FAR23 regulations and will be authorized to operate in the full range of normal instrument flight conditions. Angel's operational plan will be conservative for it will allow the

HALO aircraft to avoid significant weather by diverting to alternate municipal airports with current and forecast conditions approaching visual flight rules. Even the largest storm systems, a few hundred miles across, can be avoided, since such a distance will be readily traversed with the flight speed and fuel margin offered by the airplane.

The following factors contribute to Angel's expectation of achieving high operating reliability with HALO aircraft:

- Lessons over decades with airborne military transport and express freight aircraft operations will be applied.
- Angel will utilize fleet redundancy to ensure continuity of service in anticipation of worst case failure scenarios.
- HALO aircraft fleet operations will be steadily improved as the aircraft design is refined and as operating behaviors and performance are logged, analyzed, and utilized to upgrade choices of components and maintenance activities.
- HALO aircraft incorporate redundant mission critical systems.
- The pilots will provide real-time awareness of aircraft and communications systems through highly-responsive graphical user interfaces.

5. MARKETS & SERVICES

Angel is implementing a two-part domestic marketing strategy coupled with an international Joint Venture strategy to rapidly deploy HALO Networks worldwide.

5.1 Domestic- Phase 1

Angel is pursuing relationships with those customers who desire a ubiquitous network offering a first-mile/last-mile broadband solution independent of the legacy networks, an agile, routinely upgraded, service mix, and a wireless solution which can be quickly installed with minimum fixed capital expenditures. Angel's near-term customers are likely to be the "carriers". As a "Carrier's Carrier", Angel may serve two types of customers: (1) those who need first mile/last mile services to highspeed "backbones" or long-line carriers to better serve their customers; and (2) millimeter-wave licensed spectrum holders who want to increase their subscriber base.

5.2 Domestic- Phase 2

Angel will continue as a "Carriers Carrier", and will begin to directly serve diverse businesses. As new cities are opened by Angel, the Company may wholesale a large fraction of its capacity to Carriers and reserve the remainder for Angel-

branded broadband services. Angel anticipates offering a mix of digital information services to business customers. Angel plans to reach its customers through a variety of sales channels, including strategic marketing relationships, reciprocal resale agreements, and direct and indirect sales efforts. Examples of Angel's customers include the following:

County & City Government: Angel can provide County and City Governments with an agile broadband network to provide Fire, Police, Emergency Management, and Public Works departments with backup communications to the terrestrial networks, offering a "disaster resistant" solution. The HALO Network gives "portable" broadband connectivity to these departments.

Internet Service Providers: Regional ISPs wanting to provide business customers high speed connections independent of RBOCs and Cable companies.

Healthcare Industry: Hospitals who wish to extend their network to multiple buildings, physician's offices, Human Relations offices, and to convalescing patients at home.

Manufacturing Industry: Large corporations wishing to interconnect their office buildings, manufacturing plants, sub-contractors, customers, executives and consultants.

5.3 International

Outside the United States, Angel plans to establish regional joint ventures, involving partners working in concert to market, deploy, and operate HALO Networks at the major cities within their region. Deployments will be accomplished with the cooperation of national PTTs and/or the local wireless operators as required.

5.4 First Market – Los Angeles:

Los Angeles is a strong candidate to be Angel's first marketplace. Beta testing may commence midyear-2000 and, if market and regulatory conditions are right, Angel may enter full commercial operation soon thereafter. LA is an attractive candidate because it has fair weather, a large population, and mandates for increased telecommuting. After a successful deployment in LA, Angel and its partners will select other cities. The first 3 or 4 domestic HALO sites may be on the west coast, from San Diego to as far north as Seattle.

5.5 Competitive Service Pricing

Broadband connection prices vary extensively depending upon the distance of the business from the Central Office (CO) and the number of competitive access providers in the region. Typical pricing for a T1 Connection is \$300 to \$700 per month, and 10 Megabit (10 Based-T) at \$4,000 to \$8,000 per month. Angel has evaluated how to price its packet services based upon today's market and the competitive playing field in the year 2000 and beyond. The Company has assumed rapidly falling prices of these services and has

carefully evaluated its competitive position. The HALO Network throughput is expected to be increased from its initial 16 Gbps capacity to over 100 Gbps during the first five years.

6. HALO AIRCRAFT AS SATELLITE CONCENTRATORS

The HALO Network can serve as a valuable adjunct to a broadband satellite network. Compared to HALO Networks, satellite-based broadband systems suffer four limitations, namely: they are power limited; they require costly premise equipment; transmission frequencies are regulated by the ITU; and satellites do not serve high density population centers effectively because of their inherent power and distance limitations. The HALO Network can increase the throughput and the competitiveness of satellite systems.

HALOs can compensate for the limitations:

- 10 They have ample power. The baseline power buss of the HALO airplane can deliver more than 40 KVA or greater, typically 10 times higher than available on satellites.
- 20 Because of the relatively short range between the HALO airplane and the user (10 to 50 miles), low-cost, low-power premise equipment can be used. Only a single antenna with slow tracking is required.
- 30 The HALO Network does not require ITU coordination. Frequencies assigned to the HALO Network are within the control of the local PTT. In the U.S., for example, the HALO Network can utilize 28 GHz and 38 GHz frequency bands licensed to terrestrial services. Since the HALO Network is a "frequency agnostic", frequencies in the 3 to 20 GHz range can be used for broadband services in other countries.
- 40 The HALO Network can serve hundreds-of-thousands of broadband subscribers on a metropolitan distance scale. As such, a HALO can serve as a "concentrator" for satellite systems, through which connections to destinations outside the Cone of Commerce can be made through HALO-to-satellite links.

HALOs offer three benefits to an integrated HALO-satellite global network:

- 10 Since HALOs are positioned high above most of the atmosphere, they can easily "see" numerous satellites, even those over oceans and deserts.
- 20 The HALO-to-satellite link can use very high carrier frequencies, e.g., those above 60 GHz offering ultra-wide bandwidths.
- 30 The HALO is the front end of the integrated system. The airplane and its communications systems can be accessed each day. The integrated network can thus be steadily evolved (even radically changed, if necessary).

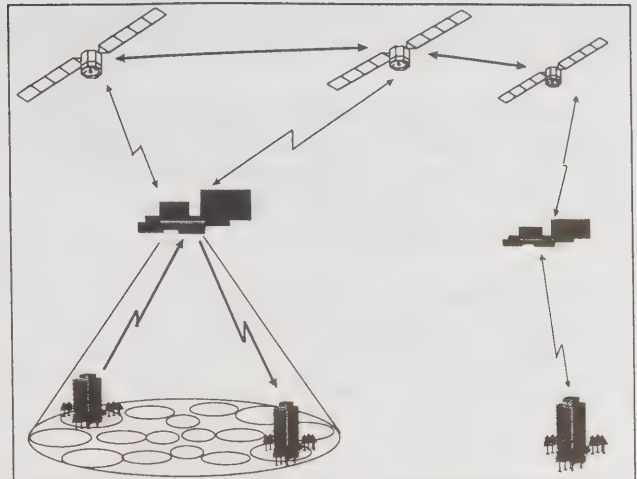


Figure #4: HALO/Satellite Hybrid Networks

7. ATTRACTIVE ATTRIBUTES OF THE HALO™ NETWORK

Given the lack of infrastructure to support the current and projected demands for broadband data communications, an intense race has begun to deploy broadband networks to satisfy businesses and consumers. These networks include satellite constellations, atmospheric alternatives to HALO aircraft, and terrestrial wired and wireless networks. The space layer includes satellites operating at geo-synchronous orbit ("GEO") and low-Earth orbit ("LEO"). The atmospheric layer includes high-altitude autonomous airships (i.e., dirigibles) and the HALO aircraft. The terrestrial layer includes wireless data transport at millimeter wave carrier frequencies, wireless local loop at the PCS frequency band, data transport over coaxial cable installed for cable television, and an assortment of digital subscriber line services ("xDSL") over twisted wire pairs installed for telephone networks.

According to multiple interviews with prospective customers, the HALO Network offers the following attractive attributes:

Ubiquity. Perhaps the most compelling attribute of the HALO Network is the ubiquitous nature of its services within the 75 mile diameter footprint. Once the HALO is flying overhead – on Day 1 of its service -- anyone within the "Cone of Commerce" can access service and communicate with anyone else within that footprint.

Upgradeability. The total capacity (throughput) of the HALO Network can be increased by swapping out the "Hub" or payload with an upgraded version.

Network Reliability/Maintenance. After each landing of the HALO aircraft, the network operations team is able to conduct a complete and thorough "check-out" of the

entire network, thus making sure that the network is operating at an optimal level of performance.

Independence from traditional carriers. The HALO Network will offer to Internet Service Providers a means of directly connecting to their customer without having to utilize Cable or Copper services, which are now beginning to compete with ISPs.

Rapid provisioning. The HALO Network can turn-on a new customer within one day, an attribute desirable to ISPs. This is best exemplified by a quote from an LA-based ISP, who said, "The Local Loop carrier everywhere is a big pain and very slow ... If we can bypass them all together that would be a great advantage. Right now if your customer asks you to install a ISDN or T1 connection, it'll really test your relationship with that customer because of the delays and difficulties involved in working with that unknown Telco."

Portable broadband services. HALO broadband services can be provided to stationary cars and trucks, for example police cars, fire trucks, and emergency management vans. Angel refers to this as "portable" broadband, of potentially high value to municipal services and to businesses such as real-estate, building/construction and transportation companies.

Easily movable/re-deployable. Because the HALO broadband connection can be so quickly provisioned, it can also be rapidly removed and reinstalled. This is of particular value to companies/organizations who wish to install temporary connections, e.g., consultants for companies, hospitals wishing to set up a connection at the home of the recovering patient, and rapidly growing businesses who are constantly juggling offices and buildings.

Disaster-Proof Infrastructure. By virtue of the airborne nature of the HALO Network, Angel is able to offer a system which is virtually disaster-proof, able to avoid the wrath which Earth quakes, tornadoes and hurricanes levy on more traditional terrestrial networks.

Angel has several parameters to vary in order to enhance the performance and to increase the capacity of the HALO Network. For example, enhancements of network capacity can be realized by increasing the number of cells in the communications pod, or the number of beam feeds per cell, or both, along with appropriate scaling of electronics. Alternatively, the bandwidth per beam spot can be widened or the encoding order increased. Performance enhancements practical at the system level, however, will be determined by trades and balances. Increasing the system capacity more than ten-fold within several years after the introduction of the first network may be realizable.

8. SUMMARY

Angel and its partners are creating a wireless broadband "super-metropolitan" area network to interconnect tens to hundreds of thousands of subscribers each at multi-megabit per second data rates. A HALO aircraft will operate above commercial airline traffic to serve as the hub of the millimeter wave wireless broadband network providing ubiquitous coverage as well as dedicated point-to-point connections. Broadband wireless services will be delivered to diverse enterprises to promote new forms of dialogue and interaction. Angel will serve two primary market segments: 'first-mile/last-mile' connections for current carriers; and broadband connections to selected businesses. The HALO Network can serve as a valuable adjunct to proposed satellite networks by serving as a "concentrator." Due to the high altitude of the HALO airplane, the HALO Network can greatly increase access to potential subscribers of terrestrial millimeter-wave services obscured by natural and man-made objects relative to towers. Angel plans to cultivate partnerships with strategic vendors, to hire engineers of superb talent, and to fostering "win-win" opportunities with satellite-based and terrestrial service providers. The HALO aircraft will commence flight testing during the summer of 1998.

REFERENCES

- 10 J. Martin and N. Colella, "Broadband Wireless Services from a HALO™ Aircraft," *Proc. of the SPIE International Symposium on Voice, Video, and Data Communications: Broadband Engineering for Multimedia Markets*, 1997.
- 20 Also refer to "The Cone of Commerce," by N. Colella and J. Martin at the same conference.

Transmission Techniques for Increased Power Efficiency in OFDM-Based Wireless Communication Systems

Rui Dinis and António Gusmão

CAPS, Instituto Superior Técnico,

Av. Rovisco Pais, 1096 Lisboa Codex, Portugal

Phone: +351 1 8419358, Fax: +351 1 8465303, e-mail: ruidinis@alfa.ist.utl.pt

Abstract

Due to their robustness against the time-dispersive effects of multipath propagation, OFDM schemes are widely considered for high rate transmission in digital wireless systems. This paper deals with power efficiency issues in OFDM-based broadband wireless systems for the mm-wave band. We consider: grossly nonlinear power amplifiers only, within conventional and two-branch transmitter structures; channel coding and interleaving in the frequency domain; a space-diversity, maximal-ratio combining scheme, and, as an alternative to conventional ARQ, similar time-diversity, packet combining schemes. A set of simulation results is presented and discussed.

Keywords: Broadband Wireless Communications, OFDM, Power Efficiency

Introduction

After having been recognized as first-choice techniques for digital broadcasting systems, OFDM techniques (Orthogonal Frequency Division Multiplexing) [1] are currently being considered for high rate transmission within wireless LAN systems working at mm-wave frequencies [2]. This is due to their robustness against severe time dispersion effects of multipath propagation, without requiring complex receiver implementations. However, OFDM modulations are often regarded as power-inefficient ones, for two main reasons: on the one hand, due to their highly variable signal envelopes, they are very prone to nonlinear distortions, namely those concerning an efficient, somewhat nonlinear, power amplification; on the other hand, BER performances are rather poor, typically approximating the BER performances of the corresponding single-carrier modulations under flat Rayleigh fading conditions.

As to the impact of the severe narrow band notches in the channel transfer function (which are inherent to the frequency-selective fading), it has been already recognized that they are no longer an obstacle to a

good transmission performance when appropriate channel coding schemes “in the frequency domain” are combined with the OFDM modulation schemes [3]. Also in recent years, several coding schemes in the frequency domain have been proposed which are able to ensure, together with their “coding gain”, a certain reduction of the signal envelope fluctuations. However, especially for a high number of subcarriers, small envelope fluctuations are only achievable if the code rate is made very small, which leads to a capacity disadvantage [4].

This paper focus on the power efficiency issues with OFDM schemes, having in mind very high bit-rate indoor systems which require mm-wave frequencies for radio transmission, high QoS (Quality of Service) levels and low-power, low-cost implementations. The paper jointly considers transmit power amplification, channel coding and diversity aspects.

Power Amplification Issues

It is well known that, when the number of subcarriers is high (say $N > 100$), conventional OFDM signals can be regarded as bandpass Gaussian noise-

like signals; their variable envelope is approximately Rayleigh-distributed, leading to a high PMEPR (Peak to Mean Envelope Power Ratio) and amplification difficulties.

A promising technique for an efficient power amplification of variable-envelope signals with a two-branch transmitter structure, proposed in [5], was considered and evaluated in [6] for the specific case of envelope-clipped OFDM signals. In this transmitter structure, two constant-envelope signals are separately amplified, without distortion, by two grossly nonlinear amplifiers, and then added to obtain the variable-envelope clipped OFDM signal. Recently [7, 8], the authors proposed a class of CEPB-OFDM schemes (Constant Envelope Paired Burst) which allows the use of a *single* grossly nonlinear power amplifier. The main idea behind these schemes is also the decomposition of variable-envelope clipped OFDM signals into a pair of signals exhibiting a constant envelope. With a CEPB-OFDM scheme, we need one nonlinear amplifier only, since the two constant-envelope bursts that correspond to a given clipped OFDM burst are just transmitted one after the other. At the receiver, we have to split the received paired burst in the two bursts and add them. The remaining of the receiver is similar to a conventional OFDM receiver.

It must be noted that practical implementations of both OFDM with a two-branch transmitter and CEPB-OFDM with a single-branch transmitter have a power efficiency advantage, when compared to conventional OFDM transmitter implementations requiring a linear power amplifier. On the other hand, these implementations don't require high oversampling factors and significant amounts of additional baseband processing effort.

Channel Coding and Diversity Issues

It is well known that OFDM schemes don't require complex receiver implementations for dealing with strongly frequency-selective Rayleigh fading conditions. However, they have poor BER performances, typically approximating the BER performances of single-carrier schemes under flat Rayleigh fading conditions, as a consequence of the almost *flat* fading at the subcarrier level. The use of space diversity techniques can help on overcoming this problem and is easy to implement at mm-wave frequencies, since an antenna spacing of about 10cm leads to almost uncorrelated CIRs (Channel Impulse Response). Obviously, the receiver structures become more complex, mainly if an MRC (Maximum Ratio

Combining) scheme and a coherent detection are intended.

Another way of improving the BER performance with OFDM schemes is by using intraburst channel coding and interleaving in the frequency domain, so as to cope with slow but strongly frequency-selective fading conditions, with no delay penalty. Convolutional codes and soft-decision Viterbi decoding are especially attractive for this purpose, combined with a conventional CRC-based error detection scheme. When an error is detected at the convolutional decoder output, a packet retransmission is requested. In a conventional ARQ scheme, the erroneous packet is discarded and the second decoding attempt relies on the second packet only; if the channel remains in a poor condition (e.g. due to a shadowing effect), the second decoding attempt will probably be unsuccessful. As an alternative to conventional ARQ, we can use "packet combining ARQ", where erroneous packets are not discarded and the different packets are combined (for example, in the MRC sense). This can be regarded as a special kind of time diversity [10]. A further performance improvement can be achieved by employing different interleavers in the several transmissions of each data block: even when the channel remains constant, a high time-diversity gain is achievable, since a given bit is affected by different fading conditions in the different transmissions.

Radio Channel Characteristics and Transmission Specifications

It is well known that time dispersion levels can be very significant in large indoor scenarios, mainly when mobile terminal (MT) antennas are assumed to have omnidirectional characteristics in the horizontal plane (delay windows above 300ns are sometimes unavoidable). Having in mind these time dispersion levels, we selected a transmission scheme with a guard interval (cyclic extension) of duration $T_G = 350$ ns in the beginning of each OFDM burst, for a burst duration of $T = 2.05\mu$ s. 324 active subcarriers and a QPSK modulation on each subcarrier were selected.

Each burst (paired burst in the CEPB-OFDM case) carries an ATM cell plus an extended header, corresponding to 57 octets. The maximum service bit rate is 160 Mbit/s for conventional OFDM and 80 Mbit/s for CEPB-OFDM. A 3-octets CRC checksum is appended to the data before the convolutional encoder. A rate- $\frac{3}{4}$ convolutional code was assumed, derived from a rate- $\frac{1}{2}$ code through puncturing [9]. This leads to 648 coded bits, of which 6 are for trellis termination (this means the number of states should be 64 or less).

The coded bits are interleaved and distributed by the 324 subcarriers.

Simulation Results and Conclusions

The performance of the proposed modulation/amplification schemes was evaluated by means of Monte-Carlo simulations. A clipping level $s_M=2.0\sigma$ was selected, with σ^2 denoting the variance of both real and imaginary parts of the OFDM complex envelope, for either conventional OFDM with a two-branch power amplification scheme or CEPB-OFDM with a conventional, single-branch, power amplification scheme; by assuming perfect receiver synchronization we achieve the same performance with both modulation/amplification schemes, provided that the two-branch transmitter is perfectly balanced [6-8].

In order to analyze the oversampling requirements with the proposed transmitter implementations, we considered FFT sizes of 512 M ($M=1, 2, 4$ and 8). It should be noted that in the present case, M is not exactly the oversampling factor; even for $M=1$ we have some oversampling, since just 324 (out of 512) active subcarriers are employed. It should be noted that the signal envelopes, at the input to the nonlinear power amplifiers, are not exactly constant for these *finite* values of M ; however, they have very small fluctuations, especially for $M=4$ and 8. Fig. 1 shows the BER performance for uncoded schemes, a power amplifier modelled as a bandpass hardlimiter, an AWGN channel and optimum receiver operation. It is clear that these small values of M are enough for an almost ideal performance. From now on, we consider $M=2$.

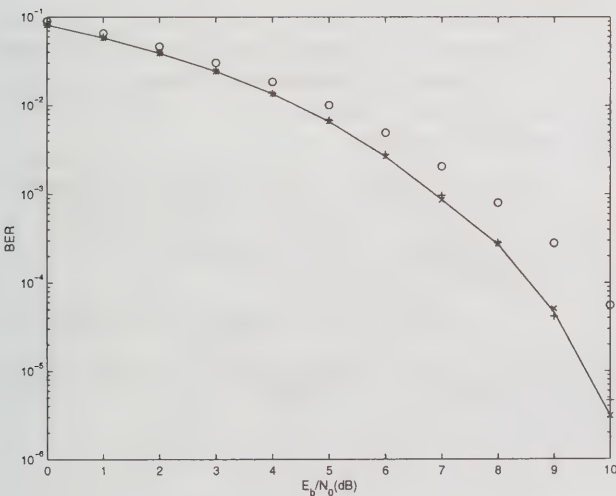


Fig. 1: BER performance for $M=1$ (o) $M=2$ (+), $M=4$ (x) and $M=8$ (solid line).

In the following, we also present a set of simulation results concerning a specific “paired-cell”, indoor scenario: a large room, with dimensions $50\text{m}\times 25\text{m}\times 4.5\text{m}$, with two base stations (BS) antennas located at the narrower lateral walls, 3m above ground, each antenna covering the opposite side of the room. A ray-tracing tool was adopted for obtaining the channel characteristics, when employing an omnidirectional antenna in the MT. A perfect channel estimation and an ideal coherent detection were always assumed, and the convolutional decoding process was supposed to take advantage of the perfect channel knowledge in an optimum way.

Fig. 2A shows the raw BER performance and the BER performance after the convolutional decoder for a 16-state convolutional code, with or without space diversity. Fig. 2B shows the probability of “decoding failure”, P_F , for the same conditions. As expected, both two-branch space diversity and channel coding lead to significant performance improvements.

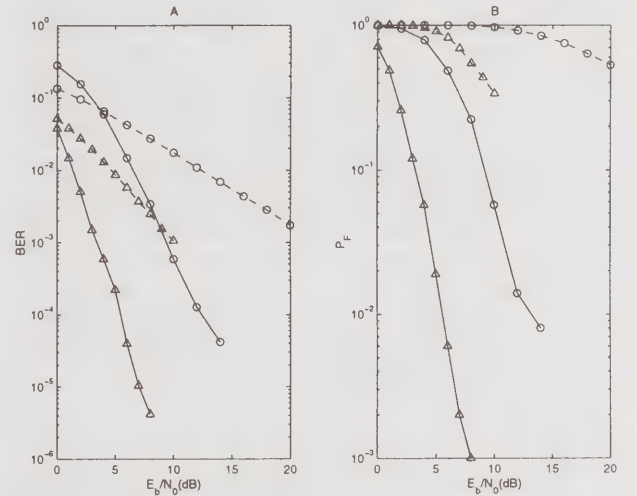


Fig. 2: BER (A) and P_F (B) before channel decoder (dashed-line) and after channel decoder (solid-line) without diversity (o) and with two-branch MRC space diversity (Δ).

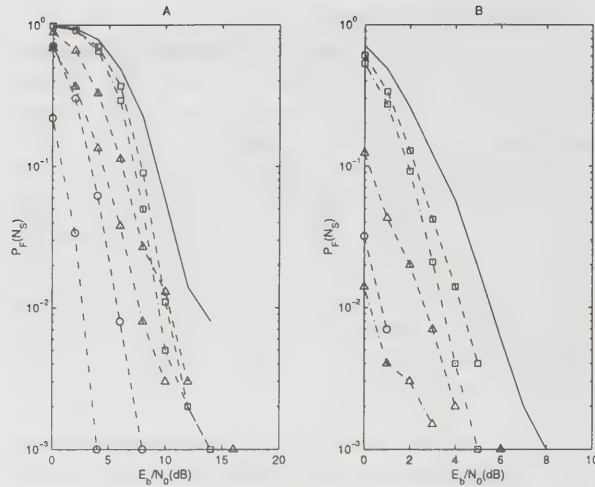


Fig. 3: $P_F(N_S)$ without diversity (A) and with MRC space diversity (B) for $N_S=1$ (solid line), $N_S=2$ (dashed line) and $N_S=3$ (dash-dotted line): (\square) for conventional ARQ, (Δ) for packet combining with constant interleaving and (\circ) for packet combining with variable interleaving.

Fig. 3 shows the evolution of the “failure probability” after N_S transmitted packets, $P_F(N_S)$, either when each unsuccessful packet is discarded (conventional ARQ) or when we keep the previous packets and combine them (packet combining ARQ); of course, $P_F(1)$ equals the probability P_F shown in Fig. 2B. Two interleaving alternatives are considered: the same interleaving for all transmissions and a different interleaving for each transmission. The channel was supposed to be the same in the several transmissions. This figure shows the advantage of packet combining, especially when a variable interleaving is employed (no failure was observed for $N_S=3$, with space diversity).

As a general conclusion, the proposed modulation/amplification schemes can give a good contribution to solve the “power problem” of multicarrier modulation, together with space diversity, frequency-domain interleaving and convolutional coding. The CEPB-OFDM scheme is particularly interesting, since a single *nonlinear* power amplifier is required at the transmitter. More specifically, the simulation results show that

- A reduced oversampling factor (4, or even 2) is adequate, for both conventional OFDM with a two-branch transmitter and CEPB-OFDM with a conventional, single-branch transmitter.

- A high-rate punctured convolutional code is enough for an acceptable performance.
- By adopting for retransmissions a different interleaving in the frequency-domain, an improved performance can be achieved, within a “packet combining ARQ” procedure.

References

- [1] - L.Cimini Jr., "Analysis and Simulation of a Digital Mobile Channel Using Orthogonal Frequency Division Multiplexing", IEEE Trans., COM-33, Jul. 1985.
- [2] - C.Ciotti and J.Borowski, "The AC006 Median Project-Overview and State of the Art", ACTS Mobile Telecomm., Summit, Granada, Spain, Nov. 1996.
- [3] - B.Le Floch, M.Alard and C.Berrou, "Coded Orthogonal Frequency Division Multiplex", Proceedings of the IEEE, Vol. 83, No. 6, Jun. 1995.
- [4] - R.D.J.van Nee, "OFDM Codes for Peak-to-Average Power Reduction and Error Correction", IEEE GLOBECOM'96, London-UK, Nov. 1996.
- [5] - D.Cox, "Linear Amplification with Nonlinear Components", IEEE Trans., COM-22, Dec. 1974.
- [6] - R.Dinis and A.Gusmão, "Performance Evaluation of OFDM Transmission with Conventional and Two-Branch Combining Power Amplification Schemes", IEEE GLOBECOM'96, London-UK, Nov. 1996.
- [7] - R.Dinis and A.Gusmão, "CEPB-OFDM: A New Technique for Multicarrier Transmission with Saturated Power Amplifiers", IEEE ICCS'96, Singapore, Nov. 1996.
- [8] - R.Dinis and A.Gusmão, "Performance Evaluation of a Multicarrier Modulation Technique Allowing Strongly Nonlinear Amplification", IEEE ICC'98, Atlanta-USA, June 1998.
- [9] - J.Cain, G.Clark and J.Geist, "Punctured Convolutional Decoding of Rate $(n-1)/n$ and Simplified Maximum Likelihood Decoding", IEEE Trans. on Comm., COM-27, Jan. 1979.
- [10] - F.Adachi, S.Ito and K.Ohno, "Performance Analysis of a Time Diversity ARQ in Land Mobile Radio", IEEE Trans. on Comm., Vol. COM-37, No. 2, Feb. 1989.

Cellular Network Design and Smart Antenna Systems

Mark Reudink
Metawave Communications
P.O. Box 97069
Redmond, WA 98073-9769
425-702-5834, 425-702-5977 (fax), markr@metawave.com

Abstract

Using switched narrow-beam antenna systems allows a frequency planner greater flexibility in designing a cellular system. In addition to flexible sector sizes, more precise traffic data is available. By studying the amount of time particular beams or antennas are being used, the probability of having an unacceptable interference condition can be determined. As a result of having improved interference reduction and more detailed system information, aggressive frequency reuse is possible. Field results confirm theoretical calculations and show that it is possible to increase system capacity by having aggressive frequency reuse without degrading network switch statistics.

Key Words: Frequency planning, smart antennas, traffic analysis, trunking

Introduction

With increased cellular demand, service providers are forced to use new methods to increase system capacity while maintaining network quality. Analog spectrum is being reduced to allow for digital services even though analog traffic continues to increase, further compounding capacity problems.

Switched narrow-beam antenna systems have been shown to increase the carrier to interference ratio (C/I) by 7 to 11 dB over conventional 3-sector and omni systems, respectively. The threshold for an acceptable call in an analog network is 18 dB C/I. This value can be obtained with a conventional system using a sectored N=7 reuse pattern. Any increase in system capacity from the N=7 reuse pattern negatively impacts system quality because the C/I level falls below 18 dB.

Simulation results show that a network utilizing a smart antenna system can achieve the same system quality with a reuse of 4 that a conventional system can with a reuse of 7. This tighter frequency reuse doubles network capacity. The simulations assume a hexagonal cell layout and uniform traffic distribution.

Since cellular networks generally don't have hexagonal cell layouts or uniform traffic, the frequency planner has to be more creative in designing a cellular network than simply implementing an N=7 reuse plan. Splitting channel sets, channel borrowing, and non-standard frequency plans are common since traffic density varies significantly from sector to sector.

Using switched narrow-beam antenna systems allows a frequency planner greater flexibility in designing a cellular system. Since traffic information is available in 30-degree increments, a more accurate picture of the network can be drawn. By studying the amount of time particular beams or antennas are being used, the probability of having an unacceptable interference condition can be determined. This, in combination with the flexibility of smart antenna systems, allows the frequency planner to optimize sector sizes and channel utilization.

In an effort to increase capacity, it is desirable to use an omni trunk group instead of sectorized trunk groups. The primary reason cell sites are sectorized is to reduce interference, but this significantly decreases trunking efficiency and, therefore, the capacity of the site. Using traffic analysis, narrow-beam antenna systems enable omni trunk groups, or near omni trunk groups, to be used without impacting call quality.

The system used in this analysis consists of 12 antenna beams, each 30 degrees wide. The mobile is served by two beams on receive and a single beam on transmit. As a mobile moves through the cell the beams serving the mobile are updated periodically. Which beams serve the mobile is based on a combination of signal strength and SAT/DSAT measurements.

Setup

The two analog cell sites shown in Figure 1 are equipped with smart antenna systems. Prior to the installation of the smart antenna system, cell A was omni and cell B had two sectors. After the installation of the smart antenna system, cell A retained its omni trunking and cell B had the gamma sector covering 120 degrees (beams 7, 8, 9, and 10), and the beta sector was expanded to cover 210 degrees (beams 3 to 9). The two sectors overlapped by 90 degrees (beams 7, 8, and 9). 120 degrees (beams 11, 12, 1 and 2), of the area was left uncovered (as it was prior to having smart antennas) due to geographic reasons. The sectors were overlapped to increase capacity by increasing the trunking efficiency.

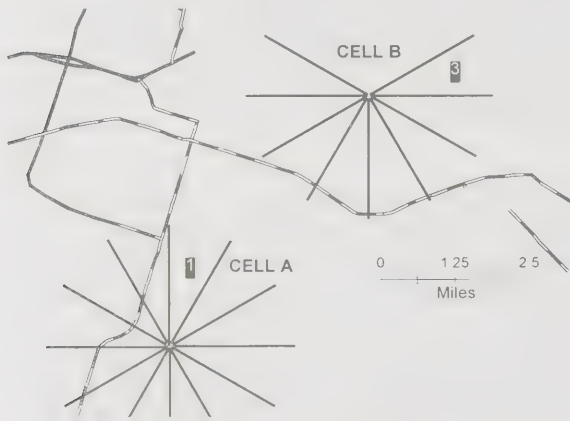


Figure 1 - Cell Locations

In order to increase capacity, it was desirable to use the same channels at the two cell sites. Drive testing and prediction plots showed that this was not feasible without smart antenna systems. In order to determine if the system quality would be compromised by having the same channels at the two sites, traffic analysis, theoretical analysis, and drive testing with the smart antenna systems were performed.

Traffic Analysis

With a uniform traffic distribution, the probability that a certain channel will be using a specific beam is $1/12$ in an omni cell site and $1/4$ in a sectorized cell site. However, uniform traffic is rare. If the majority of the traffic in cell A used beam 2 and the majority of the traffic used beam 7 in cell B, then frequencies could not be reused with the existing trunking configurations without impacting system quality. Since this configuration would be the equivalent of the existing system, therefore, the traffic distribution had to be studied. To determine the actual traffic distribution, per beam usage statistics were collected with the smart antenna system.

Table 1 shows the beam usage for Cell A and Table 2 shows the beam usage for the "Beta" sector of cell B.

Beam Number	Percent Usage
1	2.5%
2	6.0%
3	3.9%
4	3.0%
5	3.7%
6	4.0%
7	4.8%
8	5.2%
9	18.9%
10	22.0%
11	14.3%
12	11.6%

Table 1. Cell A Beam Usage

Beam Number	Percent Usage
3	15.6%
4	10.0%
5	12.3%
6	23.5%
7	23.2%
8	13.4%
9	2.0%

Table 2. Cell B Beta Sector Beam Usage

The orientation of the beams shown in Figure 1 indicates that the greatest probability of interaction between the cells is between Beam 2 at cell A and Beam 7 on cell B.

Theoretical Analysis

The theoretical analysis that determines if the same channels can be used at cell A and cell B has two parts. First, the probability of having the same channel using beam 2 of cell A and beam 7 of cell B simultaneously must be determined. Second, the probability that the mobile has a C/I less than 18 dB when this situation occurs must be determined.

The probability of having the same channel being used by both cells on the "worst case" beams (i.e. beams 2 and 7), a glare condition, depends upon how often the beams are being used and the probability the channels are on the air.

Probability of Glare =

$$(Usage of Beam 2 in Cell A) (Usage of Beam 7 in Cell B) \times (C_activity(Cell A) \times C_activity(Cell B))$$

where C_activity is the channel activity generally considered to be 80 % during busy hour.

For an unacceptable call to take place,

$$P_{Rx}(at Mobile A from Cell A) - P_{Rx}(at Mobile A from Cell B) < 18dB$$

The power received can be calculated using the Hata model.

$$P_{Rx} = P_{Tx} - C - S \times \log_{10} d + F$$

The intercept, C, is the value of the path loss 1 km from the cell and S is the slope of the path loss. The fading term, F, is a log-normal variable with zero mean and usually a standard deviation of 8 dB. The slope is usually taken between 30 and 40 dB/decade. The 1 km intercept is dependent upon the type of environment; rural, suburban, urban.

If it is assumed that Cell A Tx power = Cell B Tx power, then an acceptable call results when:

$$S \{ \log_{10}(d_1) - \log_{10}(d_2) \} + F_1 + F_2 > 18dB$$

Since all of the fading terms are log-normally distributed with the same mean and standard deviation, they can be combined into a single variable with a new standard

deviation equal to: $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$.

If the average distance from the cell to the mobile is 0.5, then $d_1=0.5$ and for an N=4 system $d_2 =$

$$\sqrt{3N} - 0.5 = 3.0.$$

Since the fading is a Gaussian random variable, the probability that the fading will cause the C/I to be less than 18 dB can be found in probability tables. Table 3 shows the probability of a C/I < 18 dB occurring for different reuse and path loss slope conditions.

Reuse	30 dB/dec	35 dB/dec	40 dB/dec
N = 4	31.0%	21.2%	15.7%
N = 7	21.2%	11.5%	6.7%

Table 3 - Probability of Interference

The probability having an unacceptable call decreases when a smart antenna system is used, since beam usage must be taken into account.

$$Probability of Poor Call = (Probability of Glare) \times Probability of Poor Call * 2 * C_activity$$

The probability of having a poor call with smart antenna systems in the system shown in Figure 1 is compared to the probability of having a poor call with the standard system in Table 4.

System	N=4, 30 dB/dec	N=4, 35 dB/dec
Smart Antenna	0.3%	0.2%
Conventional	19.8%	13.6%

Table 4 - Probability of a Unacceptable Call

Drive Testing

The quality of the system with co-channel interference was tested by transmitting the same channel at both sites while the test vehicle drove in locations with the highest possibility of interference. The tests revealed that the quality was marginal in those areas, demonstrating that this type of frequency reuse was not possible with the conventional system.

Tight Re-use Results

In order to test the system, the same five channel were used at the two sites and their performance was compared to the rest of the channels at the site over a two-week time period. Table 5 lists the results.

	Usage Time per Channel	Uplink RF loss/ Usage Time
Cell A co-channels	539	0.2%
Cell A all others	536	0.4%
Cell B co-channels	1474	0.3%
Cell B all others	1157	0.7%

Table 5 - Co-channel switch statistics

The results indicate that by studying traffic patterns and using smart antenna systems aggressive frequency reuse is realizable.

The performance quality shown by the switch statistics shown Table 5 was also confirmed with drive testing. The drive testing in the worst case area did not show any decrease in call quality. The extensive drive testing along with the switch statistics results show that having tight frequency reuse with smart antenna systems is possible without degrading network quality.

Frequency Planning Options

If the beam usage is such that an unacceptable number of poor calls are predicted, then it may be desirable to redefine the frequency plan. If Figure 1 is used as an example, then the sectors could be reconfigured such that beams 1, 2, and 3 of Cell A don't use the same channel set as beams 7, 8, and 9 of Cell B, thereby creating an interference notch. The flexible sectorization options provided by smart antennas allow operators to configure sectors that are customized for local traffic and terrain conditions. The frequency plan for the two sites could be defined as shown below. Of course, the frequency plan is dependent upon the traffic load at each site.

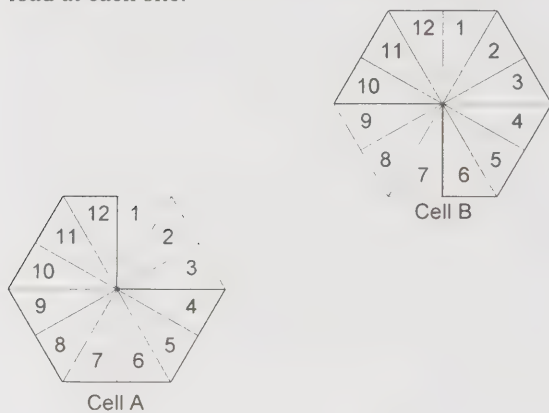


Figure 2 - Near Omni trunking with Interference Notch

Another method for increasing capacity is to combine channel from two sectors if the frequency plan allows it. The ability to reduce sectorization increases cell site capacity. A three-sector system with 15 channels per sector has a trunking efficiency of 27.0 Erlangs, 2 % GOS, Erlang B. When two sectors are combined, the trunking efficiency increases by 13% to 30.94 Erlangs.

A three-sector site was converted to two sectors with the results listed in Table 6.

System Setup	Seizures Denied (Service Busy)	% Handoff Completions	Seizures Assigned
Three Sector System	169.56	96.54	8652.89
Two Sector System	82.36	96.45	9262.27

Table 6. Sector Combining

This table shows that even though traffic increased the blocking decreased resulting in improved network service.

Summary

Narrow-beam antenna systems reduce forward and reverse link interference and allow a frequency planner flexibility in designing a network. As a result operators can use smart antennas to increase network capacity and quality. It was shown that the theoretical advantage of a narrow-beam antenna system matches the results from the field.

References

- S.C Swales, M. A. Beach, D. J. Edwards, and J. P. McGeehan, "The performance enhancement of multi-beam adaptive base-station antennas for cellular land mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 56-67, Feb. 1990.
- Y. Li, M. Feuerstein, and D. O. Reudink, "Performance Evaluation of a Cellular Base Station Multibeam Antenna," *IEEE Trans. Veh. Technol.*, vol 46, No. 1, pp. 1-9, Feb. 1997.
- M. Reudink, "Cellular System Improvements Using a Narrow-beam Antenna System," *7th Virginia Tech/MPRG Symposium on Wireless Personal Communications*, pp. 7-1 to 7-15, June 11-13, 1997.
- T. Matsumoto, S. Nishioka, and D. Hodder, "Beam-Selection Performance Analysis of a switched multibeam antenna system in mobile communications environments," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 10-20, Feb. 1997.
- J.S. Engel, "The effects of co-channel interference on the parameters of a small cell mobile telephone system," *IEEE Trans. Veh. Technol.*, vol. 18, pp. 110-116, Nov. 1969.
- M. J. Ho, G. L. Stuber, and M. D. Austin, "Performance of Switched-Beam Smart Antennas for Cellular Radio System," *IEEE Trans. Veh. Technol.*, vol. 47, No. 1, pp. 10-19, Feb. 1998.
- D. O. Reudink, D. Jones, S. Meredith, and M. Reudink, "Narrow beam switched antenna experiment," *1st Workshop on Smart Antennas in Wireless Mobile Communications*, Stanford Univ., CA, 1994.
- M. Reudink and D. O. Reudink, "C/I and SINAD Improvements from Using Switched Narrow-beam Antenna Systems in Analog Networks," *4th Workshop on Smart Antennas in Wireless Mobile Communications*, Stanford Univ., CA, 1997.

A New Throughput Analysis of A Novel MAC Protocol For CDMA-Based Wireless LANs

Peter E. Omiyi and Dr. Timothy O'Farrell, School of Electronics and Electrical Engineering,
The University of Leeds, Leeds, LS2 9JT, England.

Tel: (0113) 233 2098, Fax: (0113) 233 2032, email: eeneopo@electeng.leeds.ac.uk

ABSTRACT

A new analysis of throughput for a novel slotted CDMA-based medium-access-control protocol, with a channel-load sensing feature, for wireless LANs is presented. Numerical results are shown of throughput versus offered-load, and show that the proposed protocol offers considerable throughput improvement when compared with ALOHA DS/SSMA ('random-access' CDMA).

INTRODUCTION

Much of the current work on medium-access-control (MAC) for wireless, multimedia communications focuses on statistically multiplexing traffic over a single high-speed channel. e.g. the dynamic-slot-assignment (DSA) [6] protocols proposed for wireless ATM. On the other hand, CDMA-based MAC protocols which statistically multiplex traffic over multiple parallel channels and inherently offer improved access-delay performances, have received much less attention.

We present the slotted overload-signal spread-spectrum (OSSS) protocol for use on the up-link of wireless LANs (WLANs) with a cellular network topology. OSSS is required to co-ordinate communication between several users, each employing a unique pseudo-random spread-spectrum sequence, and a central station. OSSS falls into the category of slotted medium-access-control (MAC) protocols which use *channel-load* information to control CDMA channel-access [2-4], in order to enhance throughput in the presence of multiple-access-interference (MAI).

The OSSS protocol can be viewed as an extension of [2], however, OSSS employs a novel access strategy which ensures that the channel-load in each slot increases gradually,

and so OSSS can prevent the transmission of new packets in a slot when the MAI is too high, rather than aborting transmissions already on the channel as in [2]. The OSSS access strategy differs from the other methods proposed for controlled CDMA channel access in [3-4], which depend on the slot-size and the 'burstiness' of the traffic, respectively. The channel-load sensing protocol proposed in [3] relies on the use of slot-sizes much smaller than the packet size, in order for channel-load sensing to be effective in enhancing its throughput performance. Whereas the protocol in [4] relies on 'packet-reservation' for the estimation of channel-load, which is then used for channel-access control purpose.

In the OSSS protocol, when a packet is generated at the start of a slot, the terminal begins listening on the down-link channel and schedules to transmit the packet at a random instant within the access-window (AW) of the slot (see Figure 1). Now, the central access-point (AP) monitors the channel-load during the AW period, and broadcasts an overload-signal (OS) the instant it detects K or more transmissions on the channel.

If the OS is heard before a packet transmission begins, the transmission is deferred and re-scheduled for a future slot, otherwise the packet is transmitted. We

define K as the *overload-limit* and τ as the maximum network latency (round-trip propagation delay plus processing delay), and assume that all terminals experience this maximum latency. Due to the τ sec latency, in general, more than K transmissions would occasionally be allowed onto the channel. We define a slot as being *under-loaded* if it contains less than K transmissions and *over-loaded*, otherwise.

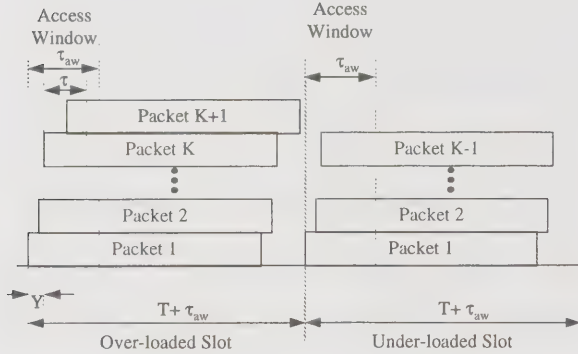


Figure 1 Channel-States of the OSSS Protocol

THROUGHPUT ANALYSIS

A single isolated picocell of a radio WLAN is considered. There is a large population of terminals and when observing the channel the distribution of the number of new and re-transmitted packets ‘arriving’ at the start of each slot is Poisson distributed with parameter GT_s and probability-distribution $\rho(k) = (GT_s)^k \exp(-GT_s) / k!$, where G is the arrival-rate and T_s is the slot-duration. All packets are of a fixed duration T seconds, with L bits per packet, and the slot-duration $T_s = T + \tau_{aw}$, where τ_{aw} is the duration of the AW. A DS/CDMA channel is employed with processing gain N , and the channel is assumed to be noiseless in the absence of MAI.

The steady-state throughput S_o is defined by $\sum_{k=1}^{\infty} k P_c(k) f(k)$ packets/slot [1], where $f(k)$

is the probability distribution (PD) of the number of transmissions in a time-slot and

$P_c(k)$ is the probability of packet transmission success given k simultaneous transmissions. For the OSSS protocol $f(k)$ is given by:

$$f(k) = P_{UL}(k) \sum_{k=0}^{K-1} \rho(k) + P_{OL}(k) \sum_{k=K}^{\infty} \rho(k) \quad (1)$$

where $\sum_{k=0}^{K-1} \rho(k)$ and $\sum_{k=K}^{\infty} \rho(k)$ are the probabilities that a slot is under-loaded and over-loaded, respectively. Also, $P_{UL}(k)$ and $P_{OL}(k)$ are the PDs of the number of transmissions in an under-loaded and over-loaded slot, respectively, where $P_{UL}(k) = 0$ for $k \geq K$ and $P_{OL}(k) = 0$ for $k < K$. $P_{UL}(k)$ is given simply as:

$$P_{UL}(k) = \rho(k) / \sum_{k=0}^{K-1} \rho(k) \quad (2)$$

Let $Y \leq y$ denote the time from the start of an over-loaded slot to the instant within the AW that the K -th transmission just begins, where $0 < y \leq \tau_{aw}$, and denote the probability-distribution function (PDF) of Y as $\Pr(Y \leq y)$. Then $\Pr(Y \leq y)$ is the probability that at least K packets are scheduled to begin transmission within the first y seconds of the AW, given that K or more packets arrive at the start of a slot. Given that the parameter of the Poisson distribution of packet-arrivals at each slot epoch is GT_s , then from the ‘splitting property’ [5] the PDs of the number of packets that are randomly scheduled to begin transmission within the first y seconds and last $\tau_{aw} - y$ seconds of the AW, comprise two independent Poisson distributions with parameters $G(y/\tau_{aw})T_s$ and $G(1 - y/\tau_{aw})T_s$, respectively. Then $\Pr(Y \leq y)$ can be derived as:

$$\Pr(Y \leq y) = \frac{\sum_{k=K}^{\infty} \{G(y/\tau_{aw})T_s\}^k \cdot \exp[-G(y/\tau_{aw})T_s] / k!}{\sum_{k=K}^{\infty} \rho(k)} \quad (3)$$

If $Y \leq \tau_{aw} - \tau$ (the duration of the AW less the latency), packet transmissions scheduled up to a maximum of τ secs after Y are not deferred. Then the PD of the number of additional packets scheduled for transmission within the interval τ is Poisson with parameter $G(\tau/\tau_{aw})T_s$. However, if $Y > \tau_{aw} - \tau$, the PD of the number of additional packets scheduled for transmission within the remaining time $\tau_{aw} - Y$ (less than τ) of the AW is Poisson with parameter $G\left(\frac{\tau_{aw} - Y}{\tau_{aw}}\right)T_s$. Then, $P_{OL}(k)$ can be derived as:

$$P_{OL}(k) = \frac{\left\{ G\left(\frac{\tau}{\tau_{aw}}\right)T_s \right\}^{k-K} \exp\left[-G\left(\frac{\tau}{\tau_{aw}}\right)T_s\right]}{(k-K)!} \times \Pr(Y \leq \tau_{aw} - \tau) \\ + \int_{\tau_{aw} - \tau}^{\tau_{aw}} \frac{\left\{ G\left(\frac{\tau_{aw} - y}{\tau_{aw}}\right)T_s \right\}^k \exp\left[-G\left(\frac{\tau_{aw} - y}{\tau_{aw}}\right)T_s\right]}{(k-K)!} \times \frac{d[\Pr(Y \leq y)]}{dy} dy \quad (4)$$

Finally, the probability $P_C(k)$ is given by $[1 - P_b(k)]^L$ [2], where $P_b(k)$ is the probability of a bit-error given k simultaneous transmissions. Assuming random codes are used, equal power reception of all signals at the AP and a noiseless channel in the absence of MAI, then an expression for $P_b(k)$ is given by [4] as $P_b(k) = Q\left(\sqrt{\frac{3N}{k-1}}\right)$.

NUMERICAL RESULTS

Numerical results are obtained using $N=64$ and $L=424$ bits (standard ATM packet size). The throughput and offered-load are normalised by the packet-transmission-time (ptt), hence the normalised throughput $S=S_o \times T/T_s$. Also, τ is assumed to be negligible compared to T on the indoor spread-spectrum channel, so a best-case normalised value $\tau/T=0.001$ is used.

Figure 2 shows plots of throughput against offered-load for OSSS with the access-window-to-latency ratio τ_{aw}/τ as a parameter and the overload-limit $K=16$. The throughput of slotted ALOHA DS/SSMA ('random-access' CDMA) [1] is plotted for comparison. This particular value of $K=16$ corresponds to the value of offered-load that yields the maximum ALOHA DS/SSMA throughput on the specified CDMA channel. OSSS offers a higher throughput than ALOHA DS/SSMA at both low and high offered-load. The larger the τ_{aw}/τ ratio the better the throughput advantage, especially at high offered-load.

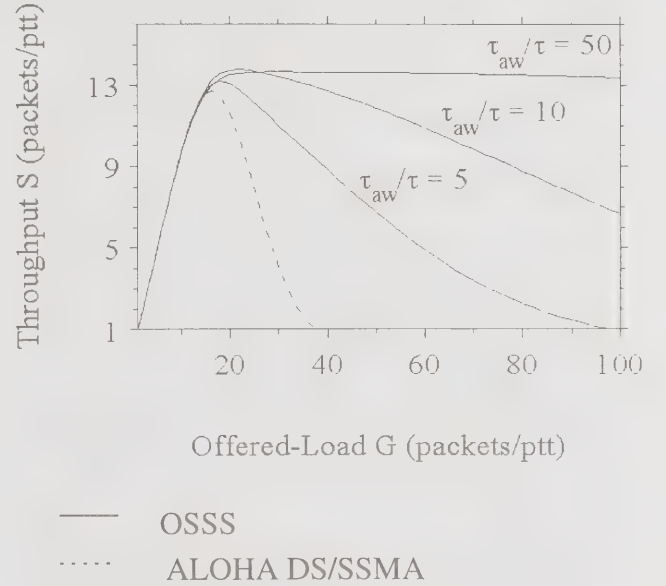


Figure 2 Throughput S versus Offered G with τ_{aw}/τ as a parameter and $K=16$.

CONCLUSIONS

A new throughput analysis is presented and used to evaluate the throughput performance of a novel OSSS protocol, proposed for WLAN (low latency) applications. OSSS is shown to offer a clear capacity advantage over slotted ALOHA DS/SSMA, enhancing the throughput at both low and high offered-load. The throughput advantage of OSSS is achieved via effective channel-load limiting in the presence of MAI.

The higher the value of the τ_{aw}/τ ratio the better the channel-load limiting capability of the OSSS protocol. Also, it can be conjectured from Figure 2 that OSSS can support a higher rate of hand-overs (as anticipated in indoor WLAN applications) than ALOHA DS/SSMA, where hand-overs will have the effect of occasionally driving the offered-traffic in a cell beyond capacity.

Unlike in [3] and [4], the throughput performance of OSSS is independent of both the slot-size and the traffic statistics, respectively, for the WLAN application considered. However, it is conceivable that 'packet-reservation' [4] can be applied to OSSS to further enhance throughput. Also, OSSS can be applied to [3] to improve its throughput performance when the slot-size is large, by providing a channel-load limiting mechanism.

The channel-access delay performance and stability of the OSSS protocol are subjects for further investigation by the authors.

REFERENCES

- [1] D. Raychaudhuri, "Performance Analysis of Random Access Packet-Switched Code Division Multiple Access Systems", IEEE Trans. on Communications, Vol. 29, No. 6, June. 1981, pp 895-901.
- [2] Dr T. O'Farrell, "Throughput Analysis of a CDMA Protocol Scheme with Channel Sensing and Overload Detection for Fibre Optic LANs", Electronics Letters, 30th September 1993, vol. 29, No 20, pp 1782-1783.
- [3] K. Toshimitsu, T. Yamazato, M. Katayama and A. Ogawa, "A Novel Spread Slotted Aloha System with Channel Load Sensing Protocol", IEEE JSAC, Vol. 12, No. 4, May 1994, pp 665-672.
- [4] A. E. Brand and A. H. Aghvami, "Performance of a Joint CDMA/PRMA Protocol for Mixed Voice/Data Transmission for Third Generation Mobile Communication", IEEE J. SAC, Vol. 14, No. 9, 1996, pp 1698-1707.
- [5] J. F. Hayes, "Modelling and Analysis of Computer Communications Networks", Premium Press, 1984.
- [6] M. J. Karol, Z. Liu and K. Y. Eng, "An Efficient Demand-Assignment Access Protocol for Wireless Packet (ATM) Networks", Wireless Networks, No 1, pp 267-279, 1995.

Comparison of FH/CDMA and DS/CDMA for Wireless Survivable Networks

Scott Chuprun and Chad S. Bergstrom

Motorola Space and Systems Technology Group
8201 E. McDowell Road, MD H1175
Scottsdale, Arizona, USA 85257-1417
602-441-4542, 602-441-2584 (fax), p15456@email.mot.com

Abstract

The future of military radio networks is changing dramatically along with commercial radio systems. Users are demanding higher bandwidths to support video, image, file transfers, and high quality voice. The methods for supplying capacity retains the centralized approach for commercial systems due to improved cell coverage, hand-off requirements, and access to wired infrastructure. The military has needs for centralized modes for minimal conflict scenarios, however, peer-to-peer modes must remain at a top priority to provide survivable communications. Timely message delivery is critical in military networks, which can have no single points of failure and must be self-organizing to be survivable.

Introduction

The results of field trials during Task Force XXI have shown the limitations of legacy radios operating in data modes to provide the needed throughput required on the battlefield. The use of MIL-STD-188-220 protocols overlaid on the SINCGARS combat net radio could not support the missions properly due to channel access delay problems, high overhead-to-data ratios, and insufficient data throughput [1]. The tests of emerging Programmable Radio technologies in these exercises demonstrated that this technology allowed unique bridging capabilities tying dissimilar networks together and providing backward compatibility.

Laboratory demonstrations of programmable radios indicate that new modulations, higher data rates, and networking function can be supported. Within the last two years, the success of computer radio development programs, government initiatives to consolidate radio programs, and the recognized need for improving network data capacity and connectivity has resulted in the Tri-Services beginning the procurement of a Joint Tactical Radio (JTR) [2,3,4]. The expectations of this type of radio platform are legacy interoperability, radio consolidation, increased data rates, support multiple access networks, and provide survivable communications.

Military Needs and Visions

Military goals for Survivable Communications are to:

- Provide significant increase in data throughputs
- Maintain interoperability with legacy systems and provide seamless bridging

- Establish survivable distributed communications networks
- Provide internet "back haul" capability with intelligent routing
- Improve video capabilities
- Demonstrate variable bandwidth systems
- Demonstrate fully distributed networks (peer-to-peer)
- Enable development of advanced portable systems
- Enhance inter-system security compatibility
- Demonstrate adaptive antenna improvements

System and Network Configuration Issues

Techniques which improve system reliability on the digital battlefield include Antenna Diversity methods, Multi-User Detection methods, and Active Interference Suppression methods. Link fidelity may be significantly improved when a variety of such techniques are used to combat terrain and mobility effects. Within a battlefield environment, the ability to provide advanced networking functions is key to the success of future voice and data radios. Within a distributed network architecture Frequency Hop (FH) or Hybrid FH / Direct Sequence (DS) transmission schemes must be optimized for access speed, throughput, and delay. These transmission method must be scaleable to meet future bandwidth needs.

A fundamental requirement for military tactical radios systems is the capability to provide intra and internet routing as well as data "back haul" functions for the digital battlefield [5]. The tactical network configurations and topologies (e.g., peer-to-peer methods, hierarchical issues, clustering) and the proposed deploy-

ment of these configurations are of critical importance in the successful design of a multiple access radio architecture. In future years, improved computer radio technology will support increased data rates and chip rates for DS signals with greater processing gain and interference survivability. These near-term gains in device technology will also improve interference avoidance and adaptive interference suppression techniques, supporting both FH and DS systems [6].

Network Multiple Access Approach

The system capacity depends on many factors such as the fading conditions, jamming, network timing accuracy, and the level of the "other user" interference. The capability of a particular signal and multiple access technique to support many simultaneous users and its complexity are significant factors in the selection of the communications system. For the case of *centralized* systems both low or moderate frequency hop (FH) rates and direct sequence (DS) systems can synchronize the individual users transmissions using the base station timing reference supported within the cellular architecture.

For a *distributed* network exposed to hostile jamming threats and unexpected channel dynamics, the use of frequency hopping systems are preferred. Frequency hopping multiple access systems do not depend on extreme power control accuracy for maintaining capacity and performance. The frequency hopping systems use the orthogonal sequences based on TOD and "network numbers" to randomly select hop frequencies and avoid the near-far problem [7]. Spectrally efficient modulation schemes using FIR filtering and power control are reducing the spectral splatter generated at the transmitter. The frequency hopped signals and networks can operate in non-contiguous bands and require less bandwidth per channel than DS wideband signals, which increases channel assignment options. Digital filtering of the DS spectrum to reduce the sidelobes within the adjacent channel is more difficult to implement. These filters require hardware FIR approaches operating at the chip rate or frequency domain techniques.

The FH system provides frequency and temporal diversity to combat the effects of frequency and time selective fading as well. Frequency hopping is ideally suited to variable data rate options, advanced interference suppression methods, and can support "10X" hop rates over SINCGARS.

Future Programmable Radio Architecture

Emerging technology advancements will enable battery-powered, man-portable radios to synthesize wide-band signals over an extreme range of operating frequencies. This section describes near-term mobile, man-pack, and hand-held military communications systems that support Frequency Hop (FH) or Direct Sequence (DS) signals. One embodiment of the programmable radio concept is shown in Figure 1, which illustrates the hardware and software platform of the SPEAKeasyII phase one architecture[8,9].

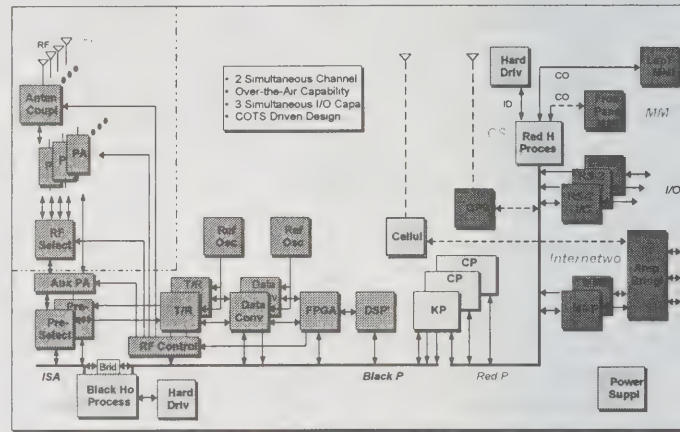


Fig. 1 - SPEAKeasyII Phase One Architecture

Programmable radios incorporate a powerful architecture for internet routing and seamless bridging between dissimilar networks. Critical subsystem elements that accommodate advanced anti-jam functions, high data rate signals, and dissimilar networks are discussed in detail below.

Programmable Radio Platform - The military is initiating the pursuit of Programmable Radio Platforms capable of meeting the communication needs described above. These computer based radios must include several transceiver subsystems to support survivable distributed communications networks.

RF Subsystem - Requires variable power control, tunable pre-selectors, excellent sensitivity and dynamic range to maintain multiple access network capacity to reduce adjacent and co-site interference at the transmission source and destination.

Frequency Synthesizer Subsystem - It must provide frequency band coverage for agility and provide a high end range at 3 - 5 GHz where large continuous blocks of spectrum exist to accommodate higher data rates and spreading bandwidths.

Modem Subsystem - DSP and data conversion devices will support the future data rate, processing gain, and interference suppression requirements for advanced anti-jam systems. The modem DSP configuration must be selected so as to accommodate FFT and FEC operations for anti-jam and high data rate network signals.

Crypto Subsystem - This system will require design influence to insure that high rate Transec keystream bits required for data encryption, frequency hop patterns, and DSPN spreading sequences are available for advanced anti-jam systems and bridging functions.

Network Router - The router has the ability to receive many types of signals from an existing network, lower or higher echelon, and transmit the information to another point in the battlefield.

Frequency Hop Network Discussion

The synchronization process for FH networks is critical to network density capability and inter-network link quality. The battlefield is composed of many user groups requiring frequency division multiple access (FDM) techniques to separate their mission traffic. Complicating this scenario, each sub-net must accommodate inter-net communications and hand-offs to provide seamless integration of forces shown in Figure 2. Frequency hopping is an excellent solution for providing network orthogonality. The timing accuracy required to allow a large number of users within the total hopset is related to the frequency hopping rate and the channel occupancy density (i.e., the rate of hop contention as a user leaves f_N , and another user arrives at f_N). As the hopping rates increase to provide improved anti-jam performance, the propagation delays become increasingly problematic relative to preservation of network orthogonality.

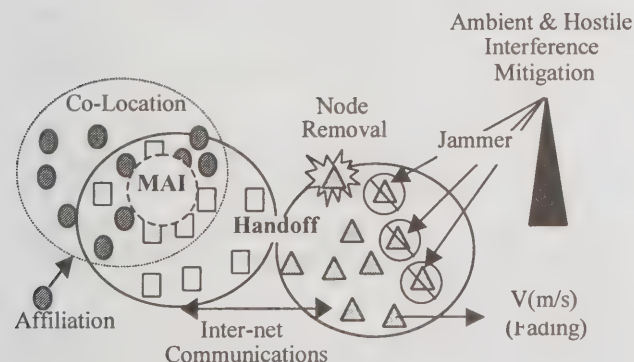


Figure 2. Frequency-Hop Multiple Access Network

Centralized Networks - For centralized networks a basestation is used to provide synchronization for the individual members. A pilot channel allows each net member to attain timing and send a request for hop pattern assignment. The basestation sends an orthogonal hop pattern assignment and a timing correction factor to align the signals being received at the basestation [10,11]. The centralized network can maintain timing more precisely than distributed networks that rely on internal TOD for peer-to-peer transmissions. It may be possible to manage time using feedback from the basestation to within a few microseconds.

Distributed Networks - The distributed network does not provide a pilot channel to establish the timing of all users in the network. Locally, timing is coordinated for all individuals within the talk group, however, there are offsets between other talk groups within the same network. This small inaccuracy is compounded by signal propagation delays, tracked by receiving users in their talk group, but creating interference for others.

The synchronization requirements for net timing can easily be attained from GPS (< 300 ns accuracy). This accuracy is adequate for system hop times in the range of 30 μs . When GPS is not available, then timing accuracy across all networks must be achieved by other means to prevent multiple access interference from degrading BER. Without GPS, it is necessary to provide a reference timing channel, or to pass Time-of-Day with a last update indication during communications.

RF front-end sensitivity is adversely impacted by the presence of adjacent networks in centralized and distributed systems. High dynamic range and coarse power control limit adjacent channel signals thus increasing the number of orthogonal mission-nets that can co-exist within a given geographical radius.

FH Network Benefits

- Same one cell frequency reuse as DS/CDMA
- Minimal intra-cell interference using orthogonal codes and proper timing
- Requires only coarse power control primarily to mitigate inter-cell interference, solves near-far problems
- Synchronization is performed at the symbol level vs. the chip level for DS
- Network can operate in non-contiguous frequency bands
- Strong jamming does not degrade entire cell or all users, instead only reduces capacity
- Avoids interference sources and frequency selective fading

Frequency Hopping Rate - Jamming threats of the adversary are limited by their ability to accurately detect the signals of interest, geometry, and processing time. The follower jammer must position itself close to the network to minimize the propagation delays and must use parallel hardware structures to reduce the signal processing tasks. These tasks include center frequency, direction-of-arrival, and time-of-arrival estimations. This ability is improving along with communications systems, however, hop rates greater than 2 to 3 Khops/second are shown to be sufficient to elude this threat [12].

To maintain network orthogonality it is important to ensure that all nets adhere to the proper TOD throughout an approximate 5 - 10 square mile range. This time-keeping function may allow one net to communicate with another (assuming the requisite Crypto capabilities) and simultaneously limit spectral dwell collisions that occurs when multiple networks are active. As networks span larger distances and the hop times approach the propagation time, orthogonality may be unachievable due to inevitable spectral collisions.

Slow frequency hopping systems can support orthogonal networks on the battlefield with less impact on performance due to co-located network timing differences. If a 2 Khop/second hop rate is assumed, then a guard time of 10 μ s can be used to allow for timing drift and to absorb degradation from multiple access traffic. Slow frequency hopping networks require much less timing accuracy than fast hopping systems, however the follower-jammer threat becomes of significant concern. Timing may be preserved via frequent transmissions or by providing a reduced hop rate synchronization channel that can be used to resolve large errors.

The main advantage of the frequency hopped system is that it provides a multiple access scheme to increase the number of users in the local area. The larger the block of hop frequencies the larger the processing gain and the number of possible orthogonal networks. In order to attain multiple orthogonal networks, each sub-net is assigned separate hop frequencies or hopping patterns. When using different hopping patterns, it is important to have very accurate timing in order to prevent two users or networks from occupying the same hop frequency.

Conclusions and Future Research

The major thrusts of advanced planning, BAA, and technology programs are to identify architectures and methods that increase the tactical data rates, provide

enhanced connectivity, and improve survivability. Our comparisons of FH/CDMA and DS/CDMA schemes examined performance issues related primarily to distributed networks and secondarily to centralized communications modes. The distributed network architecture favors using a FH/CDMA scheme for practical reasons such as drastically reduced dependence on power control, reduced timing requirements, and the ability to avoid or elude the interference. It appears that the capacity of the FH/CDMA distributed network is greater than the DS/CDMA system since it is possible to form nearly orthogonal nets. Orthogonality is not attainable for a distributed DS/CDMA which requires the network to synchronize to a small fraction of a chip time. The adaptive interference suppression methods and signal classification algorithms being developed can improve the data signal bit-error-rate and identify frequencies to avoid within hopsets.

References

- [1] US Army Signal Corps Communicator Magazine, Tactical Inter-net (EPLRS, SINCGARS SIP), Summer Issue 1997
TFXXI Lessons Learned Report, 01 July 97, Joe C. Capps
- [2] Operational Requirements Document (ORD) for the Joint Tactical Radio (JTR) 09 September 97
- [3] Future Data Radio (FDR) ORD, 19 February 97
- [4] MMITS Forum, <http://www.mmitsforum.org/>
- [5] The Warfighter Information Network master plan
<http://147.51.84.3/sysarch/win/masterplan/toc.html>.
- [6] Tradeoffs Between Frequency-Hop and Direct-Sequence Signaling for Frequency-Selective Channels, John H. Gaas, Jr. Michael B. Pursley, MILCOM '96
- [7] D. Torrieri, "Frequency Hopping and Future Army Mobile Communications", *Proceedings, ATIRP Annual Conference*, University of Maryland, 1997.
- [8] SPEAKEasyII Program, Rome Laboratory / C3BB,
<http://www.rl.af.mil/Technology/Demos/SPEAKEASY/>
- [9] The software Radio Architecture, IEEE Communications Magazine, May 1995
- [10] Frequency Hopping CDMA for Cellular Radio, Proceeding International Commsphere Symposium, Israel, December 1991
- [11] Spread Spectrum Access Methods for Wireless Communications, R. Kohno, R. Median, and L. Milstein, IEEE Communications Magazine February 1995
- [12] D. Torrieri, "Fundamental Limitations on Repeater Jamming of Frequency-Hopping Communications", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 4, May, 1989.

Dynamic Router Configuration Management for Wireless Mobile Environments

Sidarth Ambardar and Walter Lazear
The MITRE Corporation
1820 Dolly Madison Blvd.
McLean, VA 22102
(703) 695-9999, (703) 695-5213 (fax), ambardar@mitre.org

Abstract

Manual router configuration can be a tedious and time consuming task that requires significant networking expertise, and does not scale well within the emerging digitized battlefield. The dynamic nature of tactical networks, and the potential for rapid changes, requires that new router configuration files be generated and loaded into fielded routers, quickly and efficiently, to support task reorganizations. This paper describes a proof-of-concept, rule set-based, prototype tool built to convert graphical network topology representations into Cisco® router configuration files. Software "agents" tailor the configuration files and coordinate their distribution to the target routers via wireless links. The resulting ease of generating and implementing configuration files reduces dependence on highly skilled network design personnel, increases mission planning flexibility, and offers significant operational and cost benefits.

Key words: Automated router configuration, rule sets, software agents, wireless LANs, frequency hopping, spread spectrum.

1.0 Introduction

A router configuration file is a detailed set of instructions and rules that define how a router interacts with other elements of a network. Creation of this file requires significant networking expertise and attention to detail. Even small errors in a configuration file can render a routing system useless. Moreover, this task does not scale well to meet the dynamic networking needs of a tactical battlefield environment where mission requirements can change frequently, thus necessitating changes to the underlying network topologies. This potential for large-scale, unplanned, and often rapid changes to the network architecture motivates the need for an easy to use tool that can *quickly* generate and efficiently distribute the necessary configuration files. This paper summarizes a proof-of-concept effort undertaken to provide such a capability¹.

We investigated the feasibility of integrating Commercial Off The Shelf (COTS) wireless technology products with COTS software products (and custom software) to automate the generation and dissemination of router configuration files, in support of intra-network mobility and connectivity. Both of these attributes are critical to military operations.

For example, during battlefield operations the commander at a Tactical Operations Center (TOC)² may decide to move some of his command vehicles to a different location, effectively splitting the TOC. After setting up at the new location, the commander's vehicle assumes overall TOC operation, while the remaining portion packs up and later joins the command vehicles. Router configurations must be generated and disseminated to the affected routers within the TOCs, so that communications connectivity needed for mission functions is maintained.

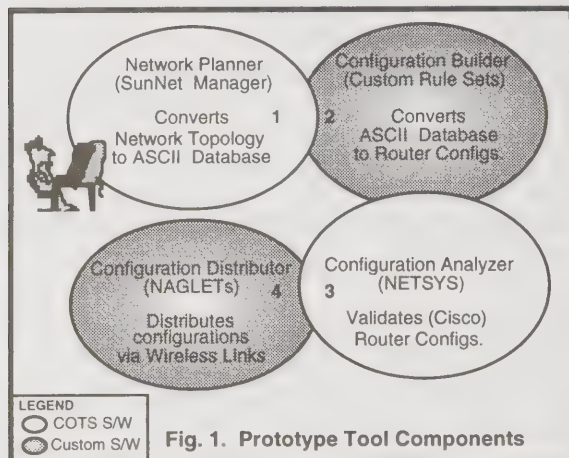
This need for a rapid configuration and reconfiguration capability is supported by observations from the Army's recently concluded Task Force XXI experiments [1]. One of the lessons learned from these experiments was that the current manual and static network planning process is not sufficiently responsive to support quick design and implementation of changing network topologies. Our effort, as described in the following sections, aimed to develop an automated tool to overcome this limitation.

¹ This work was sponsored by the Office of the Director of Information Systems for Command, Control, Communications, and Computers, US Army, Washington, DC 20103.

² A TOC is a key element of the tactical battlefield. This mobile command post is typically, formed by interconnecting vehicles via an Ethernet LAN. Information is received and transmitted internally within the TOC vehicles via this LAN, and externally via other networks and media. This facility allows the military commander and staff to integrate capabilities required to plan, coordinate, and execute mission functions.

2.0 Prototype Tool Overview

Figure 1, illustrates the four components of the prototype tool.



An overview of the functions provided by the components, and the processing sequence follows:

- 1. Network Planner:** This is a SunNet Manager™ based graphical user interface (GUI) that allows the user to draw an operational representation of a deployment scenario using a few icons (e.g., for TOC vehicles, networks, radio links) and a point and click operation. The graphical representation is stored as an ASCII database.
- 2. Configuration Builder:** This module uses the stored ASCII database and Task Force XXI network architecture specific rule sets [2] to generate candidate configuration files for each Cisco® router within the TOCs. These rule sets associate TOC vehicle attributes with actual router types (models) and configurations. They specify the configuration commands needed when a router is used to support connectivity between a TOC Local Area Network (LAN) and a wide area radio network.
- 3. Configuration Analyzer:** This module imports the generated configuration files for syntax checking and analysis. It uses a COTS product [3] to validate the generated candidate configuration files. This validation is necessary to establish high confidence in the automated configuration generation process. Without this check, any problems introduced into the configurations, such as those due to a

rule-set flaw, would go undetected until the network became operational.

- 4. Configuration Distributor:** This module stages the validated configuration files for distribution to routers over wireless links. We developed software agent³ classes to coordinate distribution of configuration files to the target routers, and to control the wireless LAN adapter units that were attached to the routers. These Java-based agents (called Network AGile appLETs (NAGLETS)) can be downloaded or instantiated remotely to perform network administration tasks (e.g., trigger router reconfiguration requests, pass alerts regarding pre-planned, pending, or priority directives, or initiate synchronized reboots) and control membership within radio networks.

The entire router configuration generation, validation, and distribution process (including intermediate processing steps) is automated. However, the user is always in control and is prompted at key decision points along the process flow. This automation eliminates configuration errors that could be introduced by hand-crafting the files. It also shortens a potentially week-long manual process into minutes⁴.

3.0 TOC Network Emulation

Figure 2 depicts a simplified TOC network layout, as set up in the laboratory. Each TOC's internal LAN was implemented using 10baseT Ethernet hubs. We used COTS wireless LAN adapter units [4] in place of military equipment to emulate the wide area radio networks and provide wireless network access to each TOC's internal LAN.

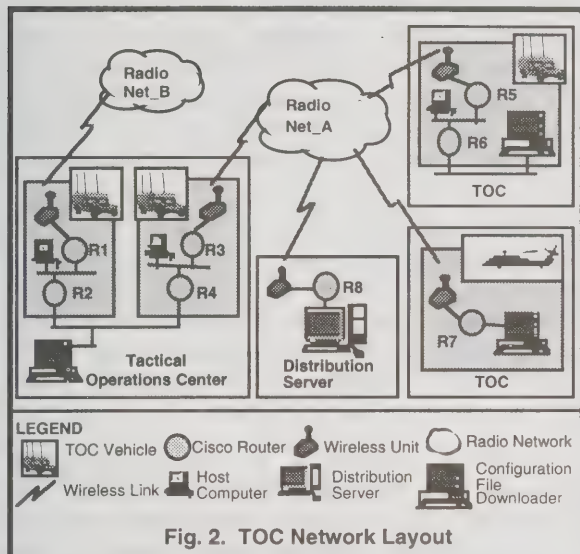
As shown in Figure 2, the wireless adapter units were connected via 10BaseT links to Cisco® routers that perform an external communications role for the TOCs. Each adapter unit⁵ has a built-in low power (250 miliwatts) frequency hopping spread spectrum

³ These agents, basically, are computer programs that are executed on remote routers to carry out operations on behalf of the Network Planner with a degree of autonomy. In so doing, the NAGLETS employ knowledge of the Network Planner's goals.

⁴ It must be emphasized that the tool does not do any traditional network planning, optimization or simulation tasks. It, basically, is an interface to allow the user to generate configuration files.

⁵ At the time this project was being conducted (September 1996-September 1997) these units were compliant with the draft D5 specification of the IEEE 802.11 subcommittee.

radio operating in the 2400 MHz to 2483.5 MHz Industrial, Scientific and Medical band. Each radio uses a pseudorandom hop pattern



to spread data transmission across time and frequency domains. Data is transmitted and received over a half-duplex wireless channel at a rate of approximately 1 Mbps. Figure 3 shows a view of the wireless LAN adapter unit.

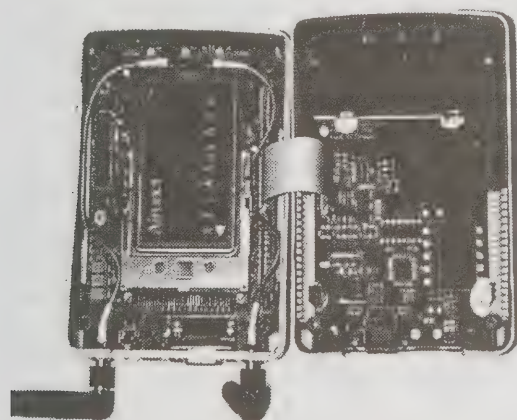
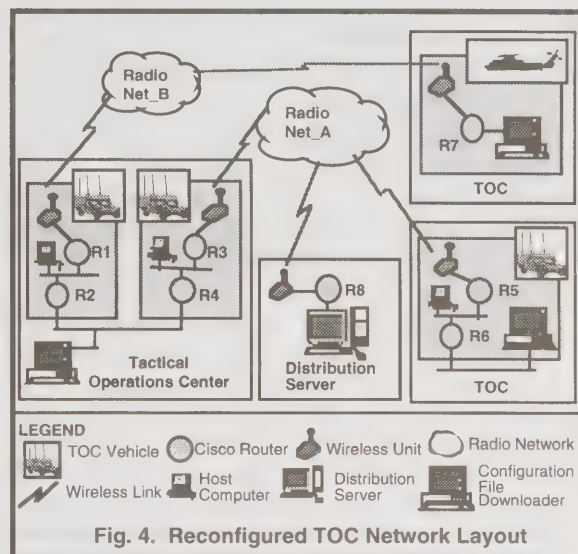


Fig. 3. Wireless LAN adapter unit

Each wireless LAN adapter unit was given an Internet Protocol address and its radio parameters were configured via a web browser. A representative radio network was created by assigning matching Extended Service Set identification values to the radios. With this setup, the wireless units supported a peer-to-peer networking relationship. A Management Information Base (MIB-1) compliant Simple Network Management Protocol (SNMP) agent allowed collection of radio performance data.

3.1 Network Reconfiguration

As indicated earlier, changes to existing network configurations are a natural part of the task reorganization process. Figure 4, depicts a reconstituted network, resulting from a typical TOC redeployment.



In this scenario, one of the TOCs (i.e. the commander's helicopter) has disconnected from Radio Net_A and has attached to Radio Net_B. To generate the configuration files needed to support this deployment, the user can either modify the baseline layout, or create a new representation on the GUI. Once the user activates the auto-generation feature, the tool determines the minimum set of routers for which configurations must be changed. Background processes generate the appropriate Cisco® configuration commands, run the validation checks, and prepare the files for distribution to the affected routers.

The tool also provides the ability to distribute only the changed portions of a configuration file (called the delta file) rather than the entire file. The delta file contains only those configuration commands that must be changed (in an baseline file) to create a new configuration file. As such, it is generally much smaller than the original file⁶. Distribution of the delta file is attractive as it creates less administrative traffic, consumes less bandwidth, and reduces latency.

Once the target router agent receives the delta file, it reconstructs the new

⁶ The size depends on the degree of change in the new topology. Typically, it varies from 5-10% of the original file size.

configuration file using the delta file and the existing configuration file. The agent then, at the specified time, sends a reboot command (via a SNMP call) to the designated router. The router loads the reconstructed file and reboots with the new configuration.

4.0 Lessons Learned

Some of the lessons learned from this effort are summarized below:

1. Automation depends on the ability to quantify roles, capabilities, and relationships. It is best applied when relationships between the TOC network components are well-defined and stable.
2. Inband distribution of configuration files requires a correctly functioning network be available all the time. The network cannot change to a new configuration until distribution is completed. Otherwise, portions of the network become unreachable. Routers that do not get the updates become "orphans" and essentially are "off the air" until new configuration(s) are manually loaded into them and they reattach to the network.
3. Although the functionality of wireless LAN products is improving, interoperability is still a long ways off. The user must carefully evaluate wireless LAN system architectures in order to maintain vendor independence and ensure interoperability across differing implementations.
4. Successful advocacy of wireless LAN technology is a challenge. Higher speed physical layers are needed to compete with technologies that provide bandwidths much higher than the currently available wireless LANs throughput of about 1Mbps.
5. The COTS configuration analysis product [3] works only for Cisco® router networks. Operation of non-Cisco routers was undefined, or at best approximated, thus reducing the product's utility.

5.0 Planned Enhancements

Enhancements being considered for the prototype tool include incorporation of:

1. Mechanisms (e.g., admission control and adaptive load balancing) for in-band distribution of network information to meet a given configuration distribution time.
2. Approaches to guarantee that "orphans" are not created. The need for inband

distribution would be eliminated if hosts and routers could automatically learn what network(s) they are attached to.

3. Security mechanisms to authenticate configuration change requests, and protect the network from malicious agents.
4. Generalized rule sets to support broader range of network architectures (e.g., Virtual LANs, and switched architectures).

6.0 Conclusion

Our investigation was conducted using technologies, some of which varied in technical maturity. The focus was to identify the potential of the technologies involved to support rapid reconfiguration capability, and not on prototype performance.

The effort produced a proof-of-concept prototype tool that automates generation and dissemination of Cisco® router configuration files in a mobile wireless environment. The tool eliminates configuration errors and helps achieve consistency in router-based networks. It reduces dependence on highly skilled network experts and reduces the costs of training and retaining such personnel.

Finally, this effort also demonstrated that wireless distribution of configuration files is possible. Careful design and implementation can deliver connectivity and mobility unattainable in a wired TOC network.

7.0 Acknowledgments

This work was conducted as a MITRE Mission Oriented Investigation and Experimentation project. We are thankful to Cecilia Hall, Dave Kocinski, and Scott Martin for contributions made to develop the software agent classes and integrate the prototype tool components.

References

1. "Task Force XXI Advanced Warfighting Experiments (16-29 March 1997); Lessons Learned", PEO C³S, Ft. Hood, Texas.
2. "TFXXI TOC Router Configuration Creation Algorithms", W. Lazear, 1997.
3. NETSYS Technologies, Inc. Enterprise/Solver™, Version 3.01.
4. Aironet Wireless Communications, Inc. Universal Client 3000E Firmware Release, Version B5.0.0.

A FDM Multiple User Software Configurable Arbitrary Format (*MUSCAF*) UHF Receiver Prototype

Ian Scott T/As Minilinx

11 Hadleigh Way

Girrawheen

Perth 6064

Western Australia

phone/fax (08) 9342 9183

email minilinx_au@usa.net or minilinx@msn.com

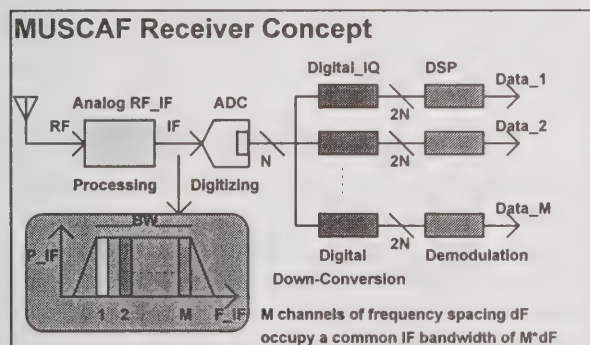
28 May 1998

Abstract

Conventional PMR style radio receivers based on Analog technology receive one frequency channel at any one time. This paper presents an alternative Digital solution which currently allows up to 16 simultaneous 12.5 kHz spaced FDM channels to be demodulated to ETSI-B Type Approval standards, and uses conventional, currently available cost effective componentry. Arbitrary Modulation Formats per Channel are supported. Technical issues up to and including the Analog To Digital Converter (ADC) interface are discussed.

1. Introduction

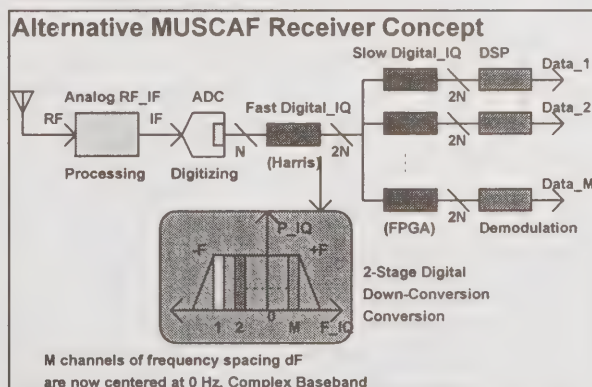
The successful design of a commercially viable MUSCAF radio requires early digitisation of the incoming RF signals. Once captured, Digital signal processing methods are used to recover signal information within a given bandwidth. Both Time and Frequency division methods can then be supported, along with arbitrary modulation formats for each user.



The Analog_RF_IF presents a processed bandwidth limited signal to a IF ADC, which digitises the composite spectrum at a relatively fast rate (e.g. 40.96 MB/s for a IF frequency of 10.7 Mhz), and produces a high speed output data bus. This data rate is then reduced to be suitable for DSP devices (e.g. 100 KB/s) using a single Digital_IQ down conversion device per channel, which also

separates I and Q signal components from the high speed data bus, and provides the IF filtering per channel in the MUSCAF receiver (each Digital_IQ "tunes" into the center of each channel and produces a lower speed IQ data stream)

If a lower overall cost per channel is required, then a two stage Digital_IQ modification allows lower cost, slower Digital IQ Down Conversions (e.g. FPGA implementation) to separate each channel.



In this concept, the first specialist Digital_IQ down-converter provides some broad "roofing" filtering, prior to channel specific filtering applied at a reduced sample rate.

2. System Considerations

The central element in the MUSCAF architecture is the Analog to Digital Converter (ADC). This component sets an upper limit to overall receiver performance due to its inherent quantization noise and distortion. Current ETSI-B Type Approval standards call for 3rd order immunity suppression of 70 dBc, and a commercial receiver will need to operate over a wide range of input signal levels, exceeding 100 dB dynamic range.

There are, in this application, 4 dominant ADC parameters to aid its selection;

Quantization Bits (N)
Sample Rate (Fs)
Maximum Input Voltage (Vpeak)
Parallel Input Resistance (Rp)

The maximum ADC input power in dBm is

$$P_{max} = 10 \cdot \log \left\{ \frac{V_{peak}^2}{2 \cdot R_p} \right\} + 30 \quad \dots(1)$$

The minimum input “aperture” power required to toggle one single bit is

$$P_{aperture} = P_{max} - 20 \cdot \log \left\{ 2^N \right\} \quad \dots(2)$$

When additional dither noise is used, this “aperture” barrier is removed, since the dither noise continually toggles bits. In this case, the actual quantization noise is a better estimate to use, and is spread uniformly from DC to one half the sample frequency Fs. In a 1 Hz bandwidth, the minimum ADC input power can be defined as

$$P_{min} = P_{max} - 20 \cdot \log \left\{ 2^N \right\} - 10 \cdot \log \left\{ \frac{F_s}{2} \right\} \quad \dots(3)$$

The resulting Noise Figure is now visible by comparing the ADC minimum input power (@ BW=1 Hz) with thermal noise (-174 dBm @ BW=1 Hz), and becomes;

$$NF_{ADC} = P_{max} - 20 \cdot \log \left\{ 2^N \right\} - 10 \cdot \log \left\{ \frac{F_s}{2} \right\} + 174 \quad \dots(4)$$

Using the 12 Bit AD9042 ADC as an example, with;

N = 12
Fs = 40.96 Mhz
Vp = 0.5 V
Rp = 250 Ohms

we predict Pmax = -3 dBm, Pmin = -148 dBm and a predicted Noise Figure of 26 dB. This final parameter is crucial in designing any successful, high

performance MUSCAF receiver, and should be minimised, since the required pre-ADC amplification must support relatively high signal levels without significant distortion.

This paper will present some of the design issues involved in the hardware implementation of a MUSCAF receiver. The system constructed interfaced an incoming UHF spectrum agile from 440-480 Mhz to a AD9042 DAC operating at 10.7 MHz with a restricted signal bandwidth of 250 KHz for convenience.

3. Some Technical Challenges - ADC RF Performance

The ADC requires some Analog processing in order to overcome its inherent weaknesses. The first weakness is a response to alias terms, and so some Bandpass filtering is mandatory. The second issue is the relatively high Noise Figures of current ADC's, compared to analog RF components such as Bipolars, FETs, MMICs, etc. (Note that Audio and RF discussions often present different noise representations, but whether we talk in nV/Hz^{0.5} or NF in each case makes no final difference.) To achieve a typical target receiver Noise Figure of 7 dB (e.g. -120 dBm sensitivity for most narrow-band FM receivers) would require a minimum pre-ADC gain of 19 dB, to overcome the ADC Noise Figure of 26 dB. This must be achieved without significant introduction of third order intermodulation products, or harmonic energy which might fall on any of the ADC/s many input alias responses. If the required radio 3rd order input intercept point is (say) +5 dBm, then the entire pre-ADC processing system must have a total output intercept point greater than +5 dBm + 19 dB = +24 dBm.

An alternative philosophy that *“the overall intermodulation of the receiver should be limited by the ADC, not the Analog processing”* suggests an even higher requirement. For 12 Bits, without dither, we would expect an intermodulation free “dynamic range” of 72 dBc. Since Pmax = -3 dBm for the AD9042 ADC, then its predicted input intercept point *would have to be* -3 dBm + 0.5 * 72 dBc = +33 dBm. When dither is applied, the intermodulation products begin to reduce due to energy spreading, and the associated intercept point is even higher.

As a result, the ADC has a bandwidth dependant intercept point, unlike other devices like Bipolars, FETs etc. When dither is used, each

discrete intermodulation product is “smeared” across a bandwidth of 0 Hz to one half the sample frequency. The bandwidth dependent reduction in each ADC intermodulation product is given by;

$$dP_{BW} \equiv 10 \cdot \text{Log} \left\{ \frac{F_Sample}{2 \cdot BW} \right\} \dots(5)$$

The “effective” ADC 3rd order input intercept point can now be established, for comparison purposes;

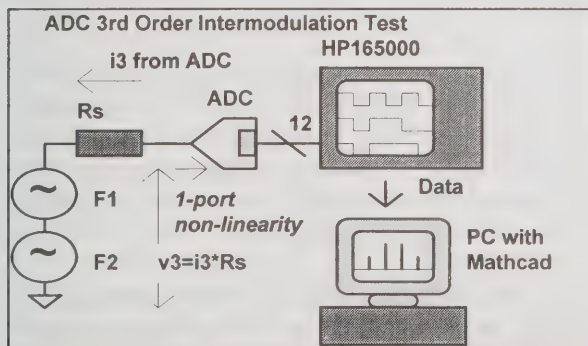
$$IP3_{ADC} \equiv P_{max} + 10 \cdot \text{Log} \left\{ 2^N \right\} + 5 \cdot \text{Log} \left\{ \frac{F_Sample}{2 \cdot BW} \right\} \dots(6)$$

In a 12.5 kHz channel plan, a signal bandwidth of 8.0 kHz is common for FM. This suggests;

BW = 8.0 kHz
F_Sample = 40.96 Mhz
N = 12
Pmax = -3.0 dBm

IP3_{ADC} = 51.7 dBm (150 Watts !)

Note that this intercept point definition only identifies an upper bound on best possible performance, and actual intermodulation performance will be worse. In particular, the ADC input impedance is not constant with signal level and can create significant 1 port non linearity. When driven with a 250 Ohm matched source, 3rd order intermodulation products at only 70 dBc were visible. These reduced ~20 dB when a 25 ohm source was used.



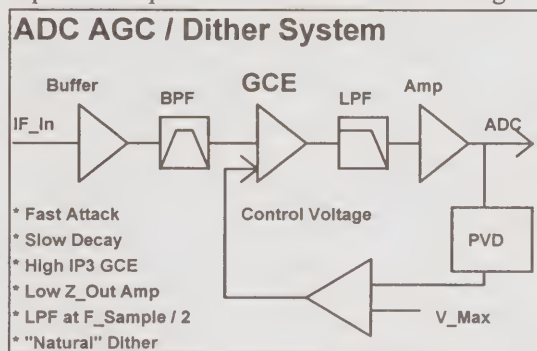
In principle, a zero source impedance should be presented to the ADC to prevent its 1-port non-linearity degrading actual intermodulation performance. The voltage amplitude of any products will be approximately proportional to Rs, so each

time Rs is halved, the corresponding 3rd order intermodulation products should drop 6 dB.

4. Some Technical Challenges - ADC AGC System

In addition, the Analog processing must account for possible ADC overload, and incorporate a very low distortion AGC system. Most AGC systems do not offer low intermodulation distortion over their gain adjustment range, and so a novel approach is required. This must have a fast gain reduction to prevent ADC overload, which will cause severe distortion and possible ADC damage.

The main system component from an IP3 perspective is the Gain Control Element (GCE) which must exhibit variable amplification (or loss) in response to a peak level derived control voltage.



In this system a Peak Voltage Detector (PVD) is used to indicate potential ADC overload and reduce system gain. A gain reduction control signal is then applied to the Gain Control Element (GCE) to prevent overload.

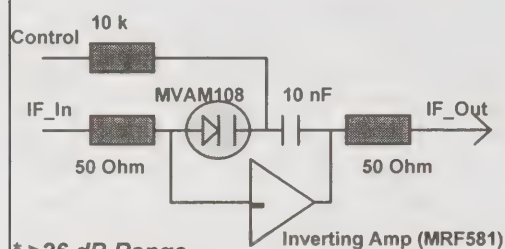
There are several choices for a GCE;

1. PIN Diode Attenuator
2. Switched Resistors
3. Variable Transconductance Device
4. DBM
5. Varicap Diode

The PIN diode seems attractive, but requires a very long carrier lifetime to exhibit low distortion at IF, especially when just conducting. Switched resistors are available in IC Formats, but exhibit poor noise figure when attenuating (bear in mind that nearby signals may activate the AGC). Variable Transconductance Devices such as Bipolar's and FET's are inherently non linear and can't be used. A DC biased DBM is feasible but suffers from a critical DC null and has an increasing noise figure with loss.

Varicap diodes however offer a wide voltage swing, acceptable gain control range and near constant NF, especially if used as a feedback element.

Varicap Diode Used As GCE



- * >26 dB Range
- * "Constant" 6 dB NF
- * $IP3(out) > 17 \text{ dBm}$
- * 50 Ohm Terminations
- * Uses AM Tuning Diode (20 - 500 pF)

Finally, some comments on "Dither". Some articles suggest that a random signal source is required at the ADC input to ensure that at least one LSB is "toggling", as a way to increase sensitivity and reduce spurs. If the ADC is used "Stand Alone" e.g. as part of a data acquisition system, then this might be necessary. However, the application of device gain and noise figure prior to the ADC will "naturally" result in Dither.

For example, Consider a 20 dB amplifier with a 6 dB noise figure placed in front of this ADC. Assume the amplifier gain is flat from DC to *one half the sample frequency* and then falls off. Since the AD9042 noise figure was 26 dB, this amplifier will present an equivalent amount of noise to the ADC, and it will naturally dither.

If a Band Pass Filter is placed in front of the ADC, this advantage will be lost.

5. More Technical Challenges-Ultra Low Distortion ADC Amplifier.

As mentioned previously, the ultimate ADC input intercept point could be as high as 51.7 dBm. This suggests that the ADC Amplifier should try to match this potential, and offer a low output impedance (< 25 Ohms).

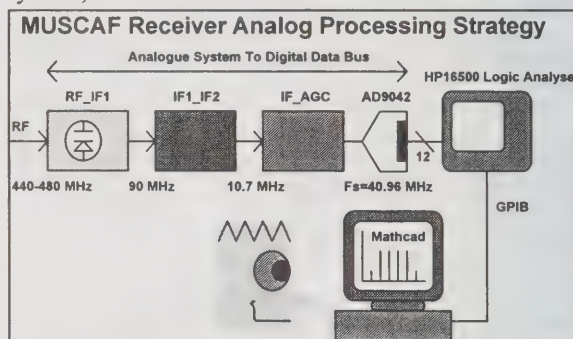
Three things are required to achieve an ultra high Output Intercept Point at IF;

1. High DC Input Power (Brute Force)
2. Negative Feedback
3. High Open Loop Gain

Current Ultra Fast OP-Amps may appear suitable, but operate in class AB, i.e. low DC power directly at the output. Various transistor arrays are available, and could be configured as a multi-stage amplifier with High Open Loop Gain, and placed in a feedback loop. An adequate solution is possible however with a single Bipolar operating with local negative feedback. (2N4427, $I_c=80 \text{ mA}$, $IP3(out) > +33 \text{ dBm}$), and is more than adequate for ETSI-B.

6. Summary

Some of the technical issues associated with the design of an example MUSCAF receiver has been presented, based on an ETSI-B compliant requirement. A practical RF-Digital MUSCAF system was constructed based on the following system;



This paper has outlined some technical issues specific to Digital capture of RF spectrums, and suggests that a commercial ETSI-B compliant MUSCAF receiver is currently feasible using conventional, obtainable componentry.

The main technical challenge is to understand the Analog To Digital Interface, and take appropriate design measures. ADC's do not exhibit conventional RF behaviour and need to be treated accordingly.

However, for comparison purposes, some familiarity with conventional terms such as Noise Figure and Intercept Point must be established. These definitions allow a A-B comparison with analog RF performance and allow us to see where the weak links in a given system design are.

This paper has presented a Noise Figure estimate for any ADC, along with various approaches for demonstrating a "Maximum Useable" Intercept Point. In addition, sub system components for a simple AGC system have been identified.

An IF Digitizing Receiver for a Combined GPS/GSM Terminal

K. Boehm*, T. Hentschel, T. Mueller*, F.Oehler***, G.Rohmer*****

* Daimler-Benz Research Ulm, FT2/HM, Wilhelm-Runge-Str. 11, 89081 Ulm, Germany

** Dresden University of Technology, Endowed Chair for Mobile Communications Systems

*** Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

Abstract

For the purpose of traffic telematics (route guidance, fleet management etc.) one needs terminals on the vehicle for the quasi-simultaneous reception of cellular phone (GSM) and satellite navigation (GPS). A conventional approach uses two complete receivers for this two tasks. We describe a cost-saving solution for a combined receiver which relies on the TDM structure of the GSM signal. Our solution is based on an bandpass-sigma-delta ADC for IF digitizing.

Key words: Traffic telematics, GPS, GSM, IF-digitizing, bandpass- $\Sigma\Delta$ -ADC, combined receiver, decimation

Introduction

The ever increasing traffic density poses problems which can be partially solved by means of modern communication systems: Route guidance, collision notification, emergency call, stolen vehicle tracking, roadside assistance, fleet- and traffic management as well as road pricing are tasks of traffic telematics. For the purpose of traffic telematics one needs terminals on the vehicle for the quasi-simultaneous reception of cellular phone(GSM) and satellite navigation(GPS). A conventional approach uses two complete separate receivers for this two tasks. In order to reduce the complexity of the total set-up and therefore achieve a reduction of volume, power and cost we propose a solution with a combined receiver which relies on the TDM structure of the GSM signal.

System Architecture

The GPS and the GSM signals can be processed in the same signal processing chain from the antennas down to the baseband. This is possible due to the time division multiplex structure of the GSM signal. The GSM needs 3/8 of the available time slots which leaves an approximately 50% duty cycle for GPS.

The set-up is depicted in the schematic of Fig. 1: The signals from the antennas (GPS=1.575GHz, GSM=.935-.960GHz) become bandpass-filtered (ceramic-filter Murata) in order to suppress strong out of band signals which could otherwise give rise to intermodulation or blocking in the preamplifier (Siemens CGY59). After the preamps variable attenuators (up to 20 dB) can be activated by the system control. Two SAW-Filters (Siemens B4678) help to suppress the image frequencies. A SPDT-switch (MACOM SW-335) serves for the time-division of the two signals. The signals are mixed down

to an IF of 44 MHz. The application of an image rejection mixer was impossible due to the large bandwidth (935MHz-1575MHz). Two separate local oscillators (LO) are switched alternately to the mixer (Motorola MC13143). An improved version may be later on realized with a single, rapid settling ($< 200 \mu s$) LO. The mixer is followed by a SAW-filter (Siemens X6959M) with a bandwidth of 2MHz in order to reduce intermodulation products in the variable gain amplifier and to prevent aliasing in the ADC. For mobile applications power consumption is a critical issue. For IF-digitizing the bandpass- $\Sigma\Delta$ -ADC has the potential of lower power consumption compared with other ADCs [Jan97]. The choice of the IF is a compromise between the requirements on the ADC (low IF wanted) and the image-rejection filter (high IF wanted). Due to the prefiltering with a bandwidth of 2 MHz an instantaneous dynamic range of 70 dB is sufficient for the ADC. The GSM specs require a (noninstantaneous) dynamic range of total 110 dB. The difference of 40 dB is supplied by the variable gain amplifier (RF2611) and the attenuators in the RF-part. In an first implementation (proof of principle) only the GPS baseband processing is realized with a PC-based evaluation board (12-channel receiver) from Plessey. From this board the RF-frontend is bypassed and the decimator output is connected directly to the correlator chip GP 201 on this board. The PC is used for control of the correlator, calculation of the coordinates, control of the VCO and the variable gain amplifier and other system control tasks.

The $\Sigma\Delta$ -modulator

Because only one IF-filter of 2 MHz (BW) is used for the combined reception of both standards (GPS, GSM), the modulator has to digitize the 200 kHz GSM-band with an

SNR of 70 dB in the presence of strong interferers in the neighbour channels. The maximum signal swing at the input is 400 mV at an $f_{IF} = 43.875$ MHz and $f_s = 19.5$ MHz. Fig. 2 shows the block diagram of the modulator architecture. It is an eighth order four stage 2-2-2-2 cascade with a fourth order noise shaping function and 1 bit quantizers in each stage. The modulator output $Y(z)$ in the Z-domain is described ideally after digital cancellation as:

$$Y(z) = z^{-8} X(z) + (1+z^{-2})^4 Q(z) \quad (1)$$

where $X(z)$ is the Z-transform of the signal and $Q(z)$ the quantization error of the last stage.

The cascaded structure was chosen to have no stability problems and less internal gain which is necessary to maximize the ratio between signal swing and supply voltage and to reduce the power consumption thereby. The internal gain increases at higher orders and therefore the input signal swing must be reduced to keep distortion low.

Track and Hold/Clock Buffer

To reduce the aperture jitter and the input impedance a fully differential, bipolar only track and hold (T/H) is used in front of the modulator. For the IF-frequency of 43,875 MHz and the sampling rate of 19,5 Msp/s the linearity of the T/H is 72 dB. To realize a signal feedthrough in the range of 70 dB with the hold capacitor of 700 fF two transistor switches are combined in series. The integrated clock buffer delivers a clock signal with a swing of 500 mV for the T/H and two differential non overlapping CMOS clocks for the loop filters, comparators and transmission gates in the DAC.

The Time Discrete Loop Filters

Fig. 3 shows the SC-implementation of the second order bandpass filters using Forward Euler loops. The integrators of the filters are fully differential single stage folded cascode amplifiers with a time continuous common mode feedback (CMFB) to keep switching noise low, as shown in Fig. 5. The unity gain frequency of this OTA is 1.7 GHz with a phase reserve of 62° . Due to different loading conditions and to reduce the power dissipation 2 versions with different transistor widths are used.

Simulation Results

In Fig. 4 the simulated noise shaping function is shown. This is the result from a full Spice simulation with a two-tone signal, each in a distance of 600 kHz from the IF center frequency and with a signal power 12 dBm below the maximum input signal power. Because of the long simulation time which is needed, this spectrum is calculated by an FFT over only 3000 points.

The total power dissipation of the modulator including the clock buffer, the track and hold and the digital cancellation filters is 150 mW on a 5 V supply.

Base-Band Processing

The base-band processing is realised by means of „off-the-shelf“ products, namely the Siemens GOLD GSM base-band processor and the Plessey GPS Correlator GP2021. Since the ADC obeys the Software Radio principle, i.e. it digitises the GSM- as well as the GPS-signals, the digitised signals have to be adapted to the constraints defined by the interface specifications of the two base-band processors. These constraints are given in Tab. 1 together with the characteristics of the signal at the output of the ADC. Since the Plessey GPS correlator is flexible with respect to its master clock rate, however the GSM base-band processor has to be clocked with 13 MHz, the overall systems master clock has been selected to be a multiple of 13 MHz. Therefore, the figures for sample rate and center frequency for the Plessey GPS correlator are based on that common clock rate, but are not the standard values suggested by Plessey. The bandwidth of the signal at the output of the ADC is determined by an analog channel filter preceding the AD conversion

	ADC Output	Siemens GSM	Plessey GPS
Sample Rate	4.875 MHz	13 MHz	4.875 MHz
Word Length	7 bit	1 bit $\Sigma\Delta$ mod.	2 bit
Bandwidth	app. 2 MHz	200 kHz	app. 2 MHz
Center	DC	DC	1.219 MHz
Frequency Resolution		10 bit	2 bit

Tab. 1: Signal Parameters given at the Output of the ADC and required at the Input of the Base-Band Processors

From Tab. 1 the functionalities of a circuitry that adapts the output signal of the ADC to the two base-band processors is obvious. These are

- Scaling, conversion from two's complement to sign-magnitude representation, and up-conversion of the signal in the GPS branch, and
- Channelisation, up-sampling, and Sigma-Delta modulation of the signal in the GSM branch.

For reasons of power-consumption the filters for channelisation are implemented using CIC-filters [Hog81] and multiplier-free FIR filters (multi-rate filter implementation). This is feasible, for mobile communications signals are to a certain degree immune to slight pass-band ripple of the filters. Moreover, the characteristics of the decimation filter at the input of the GSM base-band processor has been incorporated to the channelisation task, enabling a further simplification of

the adaptation hardware. Quantisation noise resulting from rounding in the FIR- and CIC-filters at the input of the Sigma-Delta modulator compensates for missing dithering of the 1-bit-quantiser in the modulator. Since the signals at the output of the filters still show the typical characteristics of Sigma-Delta modulated signals, the rounding operations have to be chosen and placed carefully. Another block in the GSM branch is a digital AGC, that scales the word-length of the filtered signal depending on its strength, which itself depends on the actual interferer situation at the input of the receiver.

As to the GPS-branch the ratio between sample rate and center frequency has been chosen to be a quarter. This enables a simple up-conversion by a complex multiplication of the signal with the sequences $[1\ 0\ -1\ 0]$ and $[0\ 1\ 0\ -1]$, being the digital representations of the cosine- and sine-function, respectively, at a quarter of the sample rate. All synchronisation tasks of GPS are handled within the Plessey GPS processor.

In Fig. 6 the block-diagram of the signal-adaptation circuitry is shown, indicating the multi-rate filter with the digital AGC and the low-pass Sigma-Delta modulator in the GSM branch, as well as the up-conversion, the scaling, and the signal-representation conversion in the GPS branch.

The implementation of the signal-adaptation circuitry has been realised using FPGAs. However, the intended achievement in power consumption by using multiplier-free multi-rate filters is only realisable with an ASIC implementation. Actually, the implementation of multi-

rate systems on FPGAs at maximum clock rates of approximately 40 MHz is still a problem, especially if the divided clocks are generated on the FPGAs. However, this internal clock generation is essential if scaleable and parameterisable decimation filters have to be implemented, which will be the case in the next generation of such receivers as the GSM-GPS receiver, where there will be no longer two dedicated signal paths, but the Software Radio principle will be extended to the digital domain by using general hardware, which is programmable or parameterisable by means of software.

Acknowledgement

We would like to thank Dr. J.-F. Luy (Daimler Benz) for initiating this project, helpful discussions and constant encouragement. The authors are grateful to Mr. W. Sehr and Mr. S. Urquijo (both Fraunhofer Erlangen) for their efforts in the design of the $\Sigma\Delta$ -converter

References

- [Jan97] S.A. Jantzi, R. Schreier, W.M. Snelgrove, " The design of bandpass delta-sigma ADCs" in "Delta-Sigma-Data Converters: Theory, Design and Simulation", New York, IEEE Press, 1997, ch. 9, pp. 282-308
- [Hog81] Hogenauer, E.B.: "An Economical Class of Digital Filters for Decimation and Interpolation", IEEE Trans. ASSP, Vol. 29, Apr. 1981, pp. 155-162

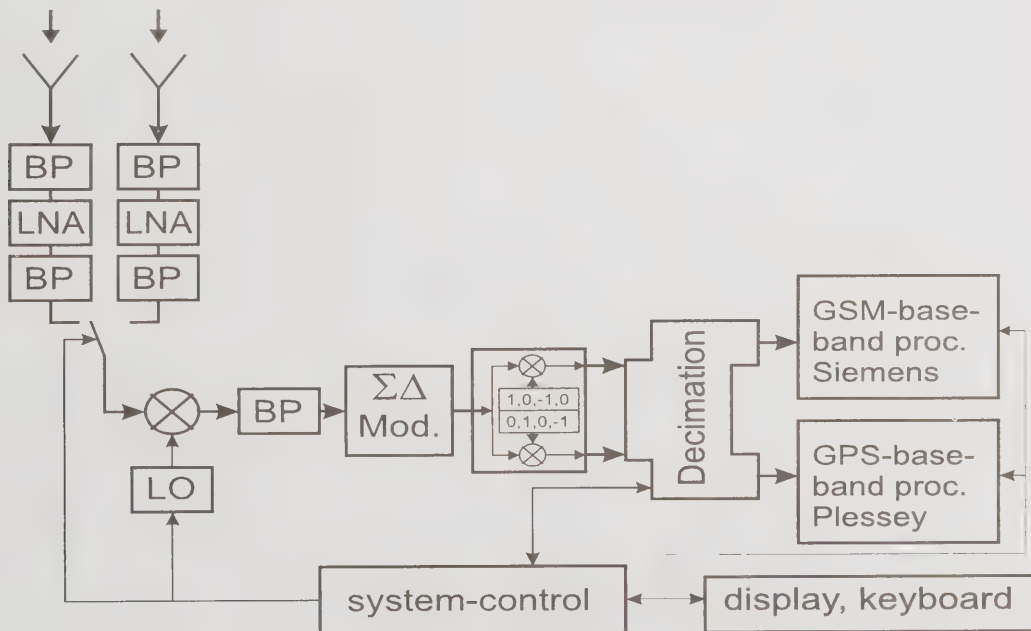


Fig. 1: System Architecture of GSM/GPS receiver

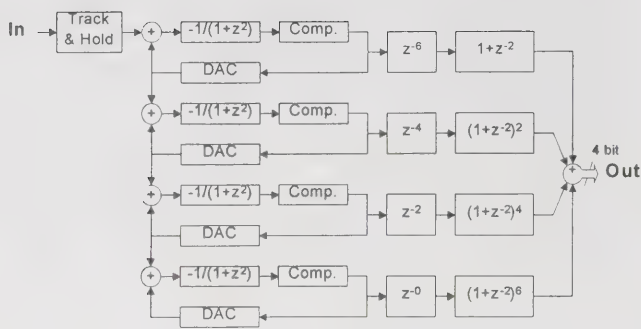


Fig. 2: Block diagram of the cascaded 8th order bandpass $\Sigma\Delta$ -modulator

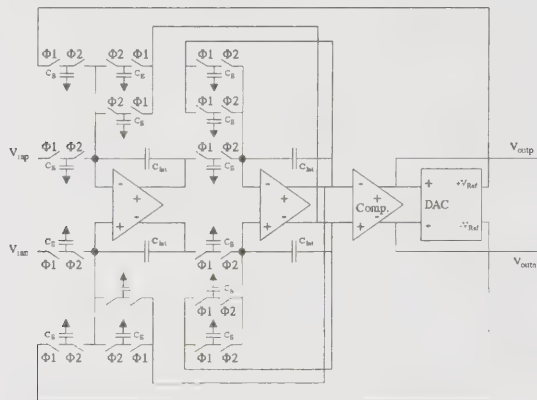


Fig. 3: SC-implementation of 2nd order modulator

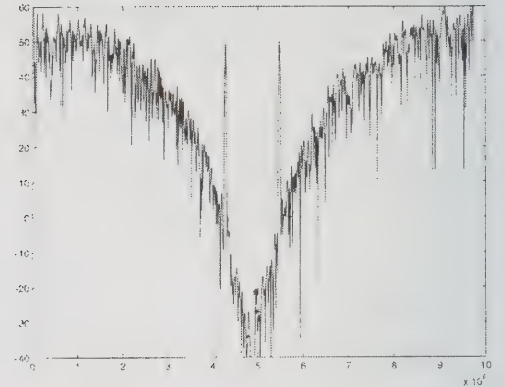


Fig. 4: Noise shaping for a two tone signal (fIF = +600 kHz/-600kHz; 100 mVpp)

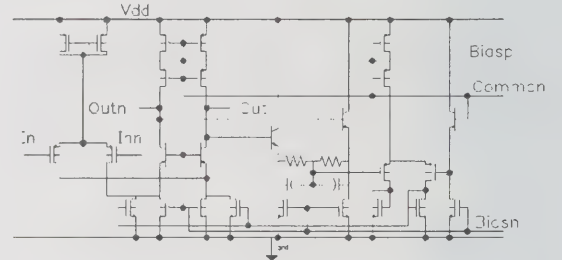


Fig. 5: Schematic of OTA

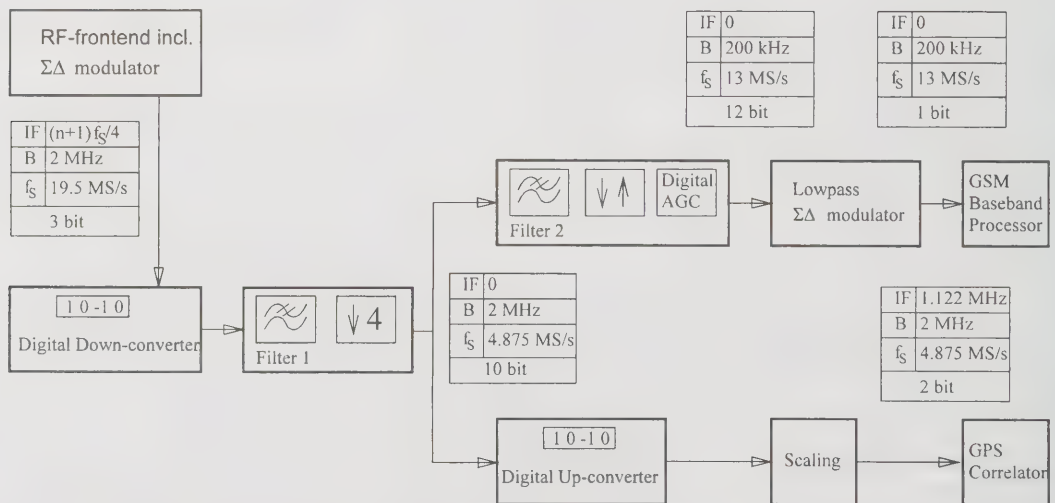


Fig. 6 : Block diagram of Signal Adaption Circuitry (decimation)

Ultra High Speed Digital Processing for Wireless Systems using Passive Microwave Logic

Bhaskar Krishnamachari, Simon Lok, Christopher Gracia, Sajan Abraham
Department of Electrical Engineering
The Cooper Union for the Advancement of Science and Art
New York, NY 10003
email: sscb@ghzdigital.com

Abstract

This paper proposes two novel high-speed digital logic families that are implemented using passive microwave circuits. The proposed logic gates can process binary information represented in two high frequency carrier modulation formats – Amplitude Shift Keying (ASK), and Binary Phase Shift Keying (BPSK). The fundamental logic gates (NOT, OR, AND) presented for both these data representations can process extremely high speed bit streams (1 - 100 Gbps). The combinational circuits formed from these passive logic gates can operate much faster than traditional electronic gates because they are not limited by the finite carrier mobility that is characteristic of semiconductors. Besides their higher operating speeds, these circuits are well suited for wireless communication systems because they process digital signals in a native high frequency transmission format and it is easy to integrate them with analog RF/microwave circuits.

Introduction

The speed of digital processing becomes more important as carrier frequencies and transmission bandwidths continue to rise in wireless communication systems. The traditional digital circuits that are used in the coding/decoding stages of such systems are limited in speed due to finite carrier mobility in switching semiconductor devices. Silicon-based devices can reach about 1.2 Gbps serial clock rates and even Gallium Arsenide based devices cannot exceed about 7 Gbps.

We propose passive microwave circuits that implement the fundamental Boolean operations NOT, OR, AND. Two logic families are presented: one operating on signals encoded using Amplitude Shift Keying (ASK) and the other operating on Binary Phase Shift Keying (BPSK) signals. These modulation formats are better for high frequency operation because the square pulses seen in traditional digital electronics are distorted at these frequencies. These gates do not use any switching semiconductor circuits because the implementation requires only passive power couplers and phase shifters. Some prior work in the area of passive microwave digital logic has been done by Kouzaev [1].

The logic gate designs presented here are not restricted in

frequency range or fabrication technology. They can be implemented for arbitrary carrier frequencies limited only by the cutoff frequencies inherent in the transmission medium.

Designs

Figure 1 shows simplified schematics of the logic gates designed and their corresponding truth tables. In the BPSK gates, the two logic states are distinguished by sinusoids ('A' and 'B') that are 180° out of phase, while in ASK gates the two logic states are distinguished by the presence of a signal ('A') or its absence ('0'). All of these gates are based on wave-interference effects. When an 'A' signal is coupled with a 'B' signal, there is destructive interference, and when two 'A' or two 'B' signals combine there is constructive interference. The power couplers shown in Figure 1 can be implemented in microstrip or stripline, using for example branch line or ring hybrid couplers. The 180° phase shifts are simply $\lambda/2$ lengths of transmission line. The gates are described in more detail below.

BPSK: The NOT gate is simply a length of transmission line corresponding to a 180° phase shift. The operation is fully reversible. The BPSK OR gate consists of two

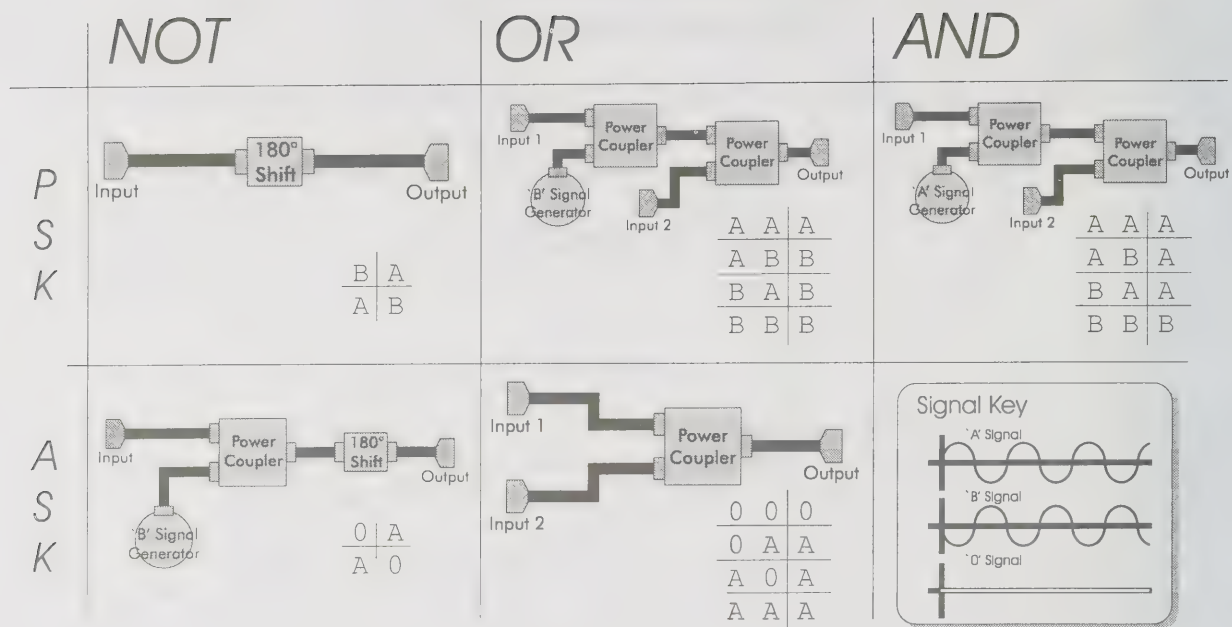


Figure 1: GHz Digital Logic Gate Designs

couplers. If input 1 is 'A', it destructively combines with the 'B' signal at the first coupler so that the output of the first coupler is always '0' which makes the output of the gate the same as input 2. If input 1 is 'B' then it constructively combines with the 'B' signal from the generator at the first coupler to generate a 'B' signal with twice the power as the output of the first coupler. In this case the output of the gate will be 'B' independent of the state of input 2. As can be seen from the designs, the BPSK AND gate is very similar in operation and is a 'dual' of the OR gate.

ASK: In the ASK NOT gate, the input is coupled to a 'B' signal generator and the output of the coupler is phase shifted by 180°. If the input is a '0' signal, the output of the coupler is a 'B' signal which then becomes an 'A' signal because of the phase shift. If the input is an 'A' signal, destructive interference takes place and the output of the coupler is '0' signal which undergoes no change in the phase shifter and emerges as the output of the gate. The ASK OR gate is simply a power coupler. The output is a '0' only if no inputs are presented to the gate and an 'A' signal if one or both inputs is an 'A' signal.

The coupler blocks shown in Figure 1 seem to be three port devices, but all passive microwave couplers are four port networks. There is dissipation of power if only one input presented to a coupler. This results in the signals attenuating as they propagate through cascaded gates. In

the worst case half of the power is lost in each gate due to this input-dependent systemic signal degradation (IDSSD). The energy required for the operations is obtained directly from the signals being processed (as power dissipation during destructive interference). In the implementation of these gates, significant power losses are a tradeoff for high speed performance.

One of the advantages of active digital logic over passive logic is that traditional digital electronic signals are regenerated after each gate. There are amplitude variations in the outputs of the passive gates (which can affect performance, particularly in ASK circuits where amplitude level discrimination becomes an issue) and phase variations that will be introduced due to fabrication errors at these high frequencies. To combat the effect of the IDSSD and other variations, active regenerating limiter/amplifier blocks may be needed after four or five cascaded stages of passive logic gates. These active regenerators can be designed to utilize existing microwave analog devices that are capable of operating at extremely high frequencies.

Simulation Results

These circuit designs have been extensively simulated over a wide range of frequencies for implementation in

microstrip and stripline. Some ASK gates have been fabricated, tested and demonstrated previously at 300 Mhz [2]. We are currently working on fabricating and testing higher frequency circuits for both ASK and BPSK. The results shown in this paper are for 10 GHz stripline circuits with a 20-mil ground plane spacing in a dielectric with ϵ_r of 9.9. The software used for design and simulation was the HP EEsof Series IV Communication Design Suite.

Figure 2 shows simulation results demonstrating the operation of a BPSK AND gate for a carrier frequency of 10 GHz. Five cycles of the carrier wave are chosen to represent one bit, so the circuit operating bit rate is 2 Gbps. The output of the OR/AND gate lags the input changes by one and one quarter carrier cycles. This is due to the physical dimensions of the circuit ($5/4 \lambda$) and represents the time delay or latency of the BPSK gates.

Figures 3 and 4 show simulation results for the ASK OR and NOT gates implemented in stripline for a carrier frequency of 10 GHz. The bit rate is again a fifth of the carrier, 2 Gbps. The output in these gates shows a time delay of half a cycle corresponding to the circuit dimensions ($\lambda/2$).

Conclusions

Designs for two novel high-speed digital logic families utilizing passive microwave circuits operating on binary information modulated using ASK and BPSK schemes have been presented. Simulations confirming the designs for operating speeds of 2 Gbps have also been shown.

There is room for significant further work in this field, particularly with regard to developing functioning multi-gate combinational circuits, and compatible memory devices. These logic families have a great potential for application in wireless communication systems because of their high operating speeds, processing of data in high frequency transmission formats and ease of integration with analog microwave circuitry.

Acknowledgements

Dr. Toby Cumberbatch, Associate Professor of Electrical Engineering at The Cooper Union, for his guidance and advice.

Dr. Aly Fathy of Sarnoff Corporation for giving us the foundations of our knowledge in microwave engineering.

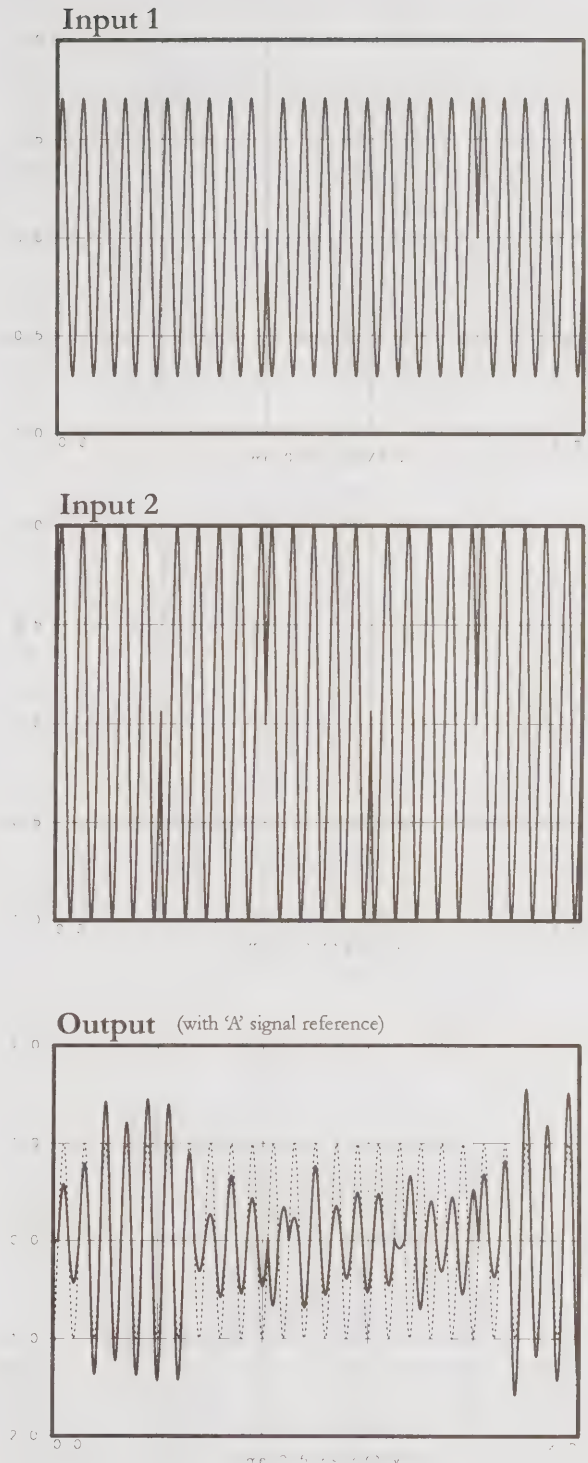


Figure 2: Simulated PSK AND Gate Results (10 GHz, 2 Gbps)

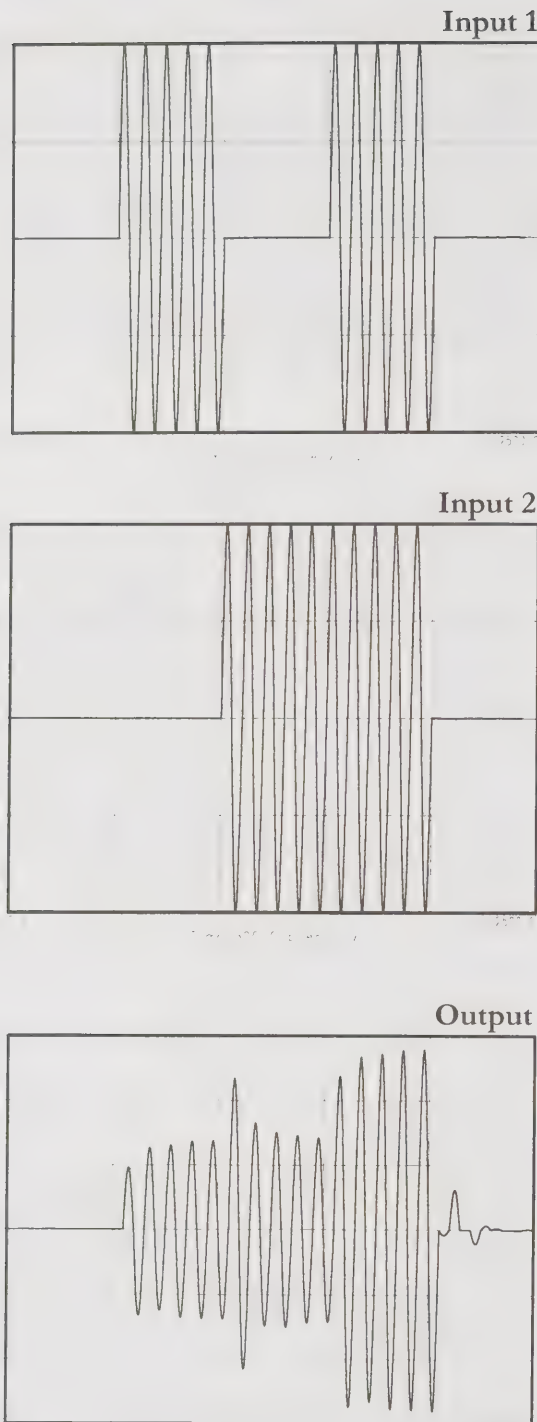


Figure 3: Simulated ASK OR Gate Results
(10 GHz, 2 Gbps)

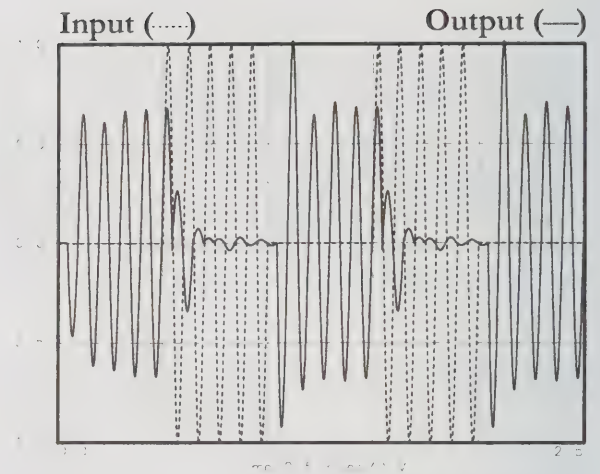


Figure 4: Simulated ASK NOT Gate Results
(10 GHz, 2 Gbps)

References

- [1] Kouzaev, G.A.; Gvozdev, V.I. "Topological pulsemodulation of field and new microwave circuits design for superspeed operating device," *1995 URSI International Symposium on Signals, Systems and Electronics*, pp. 383-384
- [2] Abraham S., Gracia, C. Krishnamachari B., Lok S. "300 MHz ASK digital logic gates in microstrip," *IEEE Sarnoff Symposium on Advances in Wired and Wireless Communications*, March 1998, pp. 173-175

A Wireless Data System Constructed of SAW-Based Receiver/Transmitter and Its Applications to Medical Cares

Kenji Matsumura^{† †}, Gen Fujita[†], Isao Shirakawa[†], and Hiroshi Inada^{†††}

[†]Dept. Information Systems Eng., Osaka University, Yamada-Oka, Suita, Osaka, 565-0871 Japan
Phone: +81(6)879-7808, Fax: +81(6)875-5902, E-mail: {matsu, fujita, sirakawa}@ise.eng.osaka-u.ac.jp

^{††}KCS Co., Ltd., Naka-Kosaka, Higashi-Osaka, Osaka, 577-0804 Japan

Phone: +81(6)722-0606, Fax: +81(6)722-7546, E-mail: matsu@k-c-s.k-c-s.co.jp

^{†††}National Cardiovascular Center Research Institute, Fujishiro-Dai, Suita, Osaka, 565-0873 Japan
Phone: +81(6)833-5012, Fax: +81(6)872-7513, E-mail: inadah@ri.ncvc.go.jp

Abstract

The present paper describes a wireless data system dedicated to medical cares by employing SAW-based receiver and transmitter. In order to realize the portability as well as to lower the power consumption, an ASIC architecture is additionally devised to implement the intermittent drive and the communication protocol, by means of which the wireless data system can admit portable monitoring facilities in medical cares, i.e. the wandering prevention of patients with senile dementia and the 24-hour observation of portable electrocardiographs.

1. Introduction

With advances of short range, unlicensed RF (radio frequency) links, different applications have been required, which include wireless security systems, remote entry systems, wireless bar code readers, wireless computer links, and so forth[1].

In order to achieve high performance and miniaturization in such wireless communications and high-speed computing applications, the SAW (surface acoustic wave) device technology has been unified with the IC technology. Thus, according to a wide variety of low-power wireless applications, such as automotive keyless entry, remote utility meter reading, wireless data links, etc., are growing rapidly, the use of SAW devices is accelerated due to the advantage of the technology to enhance reliable performance, small size, long battery life, and regulatory compliance.

To cope with conventional inductor/capacitor-stabilized super-regenerative receivers with their inherent problems of poor stability, little frequency selectivity, and poor sensitivity, the so-called SAW-stabilized super-regenerative receivers and SAW-stabilized super-heterodyne receivers have been developed, and more recently another new type of receivers, called *amplifier-sequenced hybrid* (ASH) receivers, have been conceived to address the shortcomings of previous receiver architectures in short range RF link applications[2].

By using a pair of commercially available SAW-stabilized transmitter and receiver in conjunction with an ASIC module additionally devised for intermittent drive and the communication protocol,

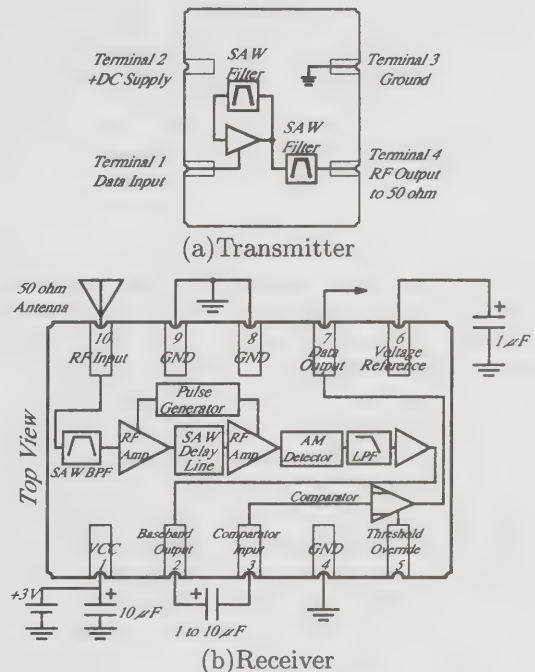


Fig. 1 Block diagram of Transmitter and Receiver.

a wireless data system can be constructed of base-stations and an IC-card type of portable terminals, which is intended to be applied dedicatedly to the monitoring for medical cares.

The present paper outlines a pair of the employed SAW-based transmitter and receiver, devises an ASIC architecture for the intermittent drive and communication protocol to be incorporated with the transmitter and receiver, and then describes a wireless data terminal with the use of these transmitter/receiver and ASIC module. The main focus is put on how to enhance the terminal performance and portability to be exploited for the wandering prevention of patients with senile dementia and the 24-hour observation of patients with portable electrocardiographs.

2. SAW-Based Transmitter and Receiver

Fig. 1 shows block diagrams of the adopted SAW-based transmitter module (HX1002) and receiver module (RX1110)[3]. The specifications for these

transmitter and receiver are as follows: (i) frequency: 303.825 MHz, (ii) bit-rate: 9,600 bps, (iii) modulation: OOK (ON/OFF keyed) modulation, and (iv) link range: 15 m.

SAW Transmitter[3]: This is a miniature transmitter module which generates OOK modulation from external signals. The carrier frequency is quartz SAW stabilized, and output harmonics are suppressed by a SAW filter. In order to implement this SAW filter, a SAW coupled resonator is employed, which has a two-pole frequency response and 3-5 dB of insertion loss, untuned, in a 50 ohm test fixture, at a center frequency of 303.825 MHz. With this scheme of using the same resonator as the output filter as well, the fundamental output power of the transmitter can be high and still provide the necessary harmonic rejection. This also allows that the output power of the transmitter can be made as high as possible in the design requiring a minimum amount of power supply current, extending battery life.

SAW Receiver[3]: Two SAW devices are employed; a low-loss coupled-resonator SAW filter performs the front-end filtering which provides fair selectivity, and a low-loss SAW delay line provides the time delay necessary to sequence the two RF amplifiers.

Referring to Fig. 1(b), the incoming signal is first selected by the SAW bandpass filter and then applied to the first RF amplifier. This amplifier is turned on by the pulse generator. The output of the amplifier is then applied to the input of a SAW delay line. As readily verified, the second RF amplifier is turned off when the first RF amplifier is on and vice versa. When the signal is emerging from the delay line, the first amplifier is turned off and the second amplifier is turned on. The output of the second amplifier is then applied to a detector circuit. Since these two amplifiers are not on at the same time, feedback from one amplifier to the other does not cause the circuit to become unstable. The delay of the delay line is chosen to obtain hundreds of samples per incoming data bit. A typical delay used is $0.5 \mu\text{s}$. The gating signal is then simply removed from the data signal with a low-pass filter following the detector.

Fig. 2 shows a diagram of the sequential gain timing[2]: (i) RF Amp Input represents the pulse on/off modulated RF signal at the input to the first RF amplifier. A data "1" is carrier on and a data "0" is carrier off. (ii) RF1 indicates the clocking of the first RF amplifier. When the pulse is high, the amplifier is turned on and vice versa. (iii) SAW Input denotes the RF pulse applied to the input of the SAW delay line. (iv) SAW Output represents the RF output of the SAW delay line. (v) RF2 indicates the clocking of the second RF amplifier.

This figure demonstrates that the data is sampled per bit at 250 kHz by the switching of the two RF amplifiers, and these two RF amplifiers are clocked

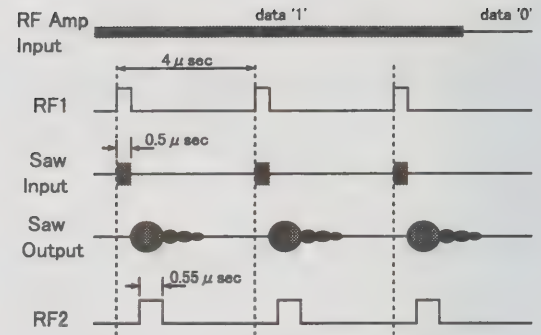


Fig. 2 Sequential gain timing.

sequentially to prevent feedback instability. Specifically, the second RF amplifier is turned on immediately after the first RF amplifier is turned off, and the "on" time of the first RF amplifier is equal to or slightly greater than the delay of the SAW delay line; hence the "on" time of the second RF amplifier is approximately 1.1 times the "on" time of the first RF amplifier, since the SAW delay line stretches the pulse time to some extent.

3. Wireless Data Terminal

Our object is to construct a wireless data system composed of base stations and an IC-card type of terminals with the use of these SAW-based transmitter and receiver, which is intended to be applied to medical cares. To attain this object, we propose a wireless data terminal ADC (Active Data Carrier) as outlined in Fig. 3. In this diagram, Selector is to switch the terminal into either the transmitting phase or the receiving phase. In order to make this terminal portable, ASIC Module is devised to be incorporated with SAW-based Transmitter and Receiver. ROM (read only memory) is for memorizing the ID (identification) of each terminal, and RAM (random access memory) is used for the transmission buffer. Electrocardiographs IF (interface) is optionally provided, to be touched on later.

The total system performance depends primarily on ASIC Module, which is in charge of not only communication protocol but also intermittent drive for enhancing the portability of the terminal. Fig. 4 shows a block diagram of this ASIC Module, which is composed of functional units; (i) Intermittent Drive Controller, (ii) ID (identification) Checker, (iii) ID Transmitter, (iv) Receiver Controller, (v) Transmitter Controller, and (vi) Electrocardiographs IF (interface), which are outlined in what follows.

[I] Intermittent Drive Controller: This unit is connected to a battery of 3V, and drives SAW Receiver and other prescribed units intermittently for 50 ms per second in the 'idling mode'. When Receiver Control detects a pulse sequence, henceforth called *ENQ Sequence* (enquiry pulse sequence), which indicates the start of the 'data receiving mode', it has to inform Intermittent Drive Controller (i) to drive all the prescribed units, i.e. to switch them from the idling mode into the data receiving mode, and (ii) to continue to drive them until Re-

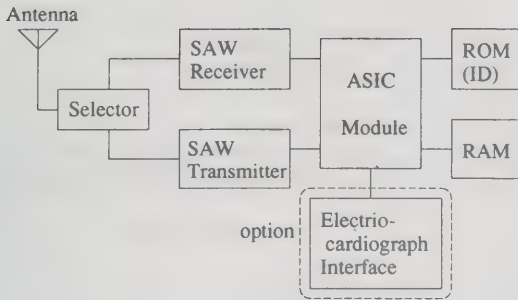


Fig. 3 Block diagram of ADC.

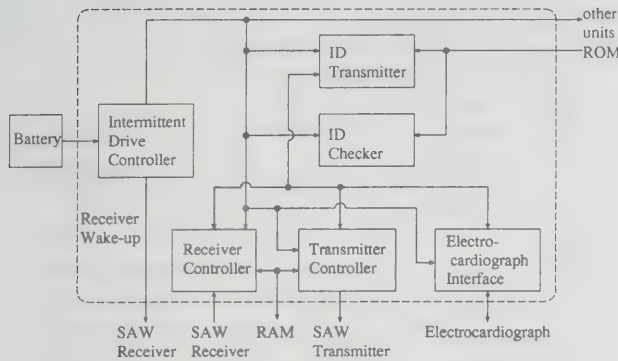


Fig. 4 Block diagram of ASIC Module.

ceiver Controller detects *EOT* (end of transmission) *Signal*, which indicates the end of the data receiving mode, and then (iii) to resume the idling mode. This means that a protocol is regulated between Receiver Controller and Intermittent Drive Controller, the details of which are to be described in [IV].

[II] **ID Checker:** With the use of an ID list registered in ROM (read only memory), this unit checks if a received ID is in the list, and then informs ID Transmitter whether or not the received ID is registered.

[III] **ID Transmitter:** This unit transfers the received ID to SAW Transmitter if it is registered.

[IV] **Receiver Controller:** First, it should be pointed out that SAW Receiver is usually unstable for 90 ms from the start of drive. Thus, as soon as this Receiver Controller detects ENQ Sequence, it informs Intermittent Drive Controller to start driving all the prescribed units so as to switch them into the data receiving mode, skips all the received pulses for the first 90 ms, then resumes reading received pulses, and as soon as it detects EOT, it informs Intermittent Drive Controller to switch all the units into the idling mode.

To make the total mechanism operate precisely, a state machine is modeled on a number of specific states, such as *ACK* (acknowledge), *NAK* (negative acknowledge), *EOT*, *time-out*, *demand for data transmission*, and on the basis of this state machine a protocol scheme is constructed as shown in Fig. 5.

In this protocol a number of typical procedures are specified. For example, while the basestation is com-

municating with a terminal, it should never issue ENQ Sequence (see lines 2-11 in Fig. 5). In addition, although the probability is very low, it may happen that two or more terminals can receive ENQ Sequence simultaneously. In this case, let each such terminal transmit its own ID to the basestation, then such ID signals cause a transmission conflict, and hence an parity error, in the basestation, and therefore let the basestation issue EOT Signal so as to start the timer of each such terminal by using random numbers (see lines 18-25 in Fig. 5).

It should be added that this Controller reads received pulses by the unit of 8-bit data plus 1-bit parity, each to undergo the parity-check.

[V] **Transmitter Controller:** This unit refines the necessary data, such as ID data, control data of ACK, NAK, EOT, etc., and optionally electrocardiograph data, into the serial form, and then transfers such serial data to SAW Transmitter.

[VI] **Electrocardiograph IF:** This unit is to read the data of the electrocardiograph optionally connected to the terminal, and then to transfer them to Transmitter Controller.

4. Applications to Medical Cares

This wireless data system can be applied to the monitoring in a variety of medical cares, mainly because a data terminal can be constructed in the form of an IC-card, and moreover, due to a precisely refined set of protocols for the intermittent drive, the power consumption of the data terminal can be lowered to 300 μ W in the idling mode and 9 mW in the data receiving mode.

Now, seeing that there occur a number of serious issues especially in the field of medical cares, we have attempted to apply this wireless data system to the monitoring in such medical cares. First, it should be pointed out that according as the rate of the population of senescent people to the total is now more and more increasing, the facilities for the wandering prevention of patients with senile dementia are of urgent necessity. Motivated by this social requirement, we exploit this wireless data system for such a medical care.

4.1 Wandering Prevention [4]

Since there occurs the tendency that the number of patients with senile dementia is growing very rapidly, there has been an increasing demand for the facilities of wandering prevention of these patients. For this purpose, our wireless data system is exploited for the monitoring of the patients in the following manner.

(i) A certain number of basestations are placed at prescribed locations inside and outside of a hospital for inmates. Each of the basestations is connected to a wireless data concentrator through RS-232C interface, and all the concentrators are interconnected to a server by Ethernet.

(ii) A data terminal is sewed into clothes of each inmate, each with a prescribed ID.

(iii) The server is used for the central monitoring, that is, it centralizes all the necessary information

about the inmates, such as the movement locus of each inmate, including his/her present location. In addition, let rank the places according to their relative urgency (for example, rank 1 := primary gates, rank 2 := doorways, rank 3 := dining room or refresh room, rank 4 := bedrooms, etc.), and the server can provide a table of names of inmates arranged according to their ranks.

4.2 Data Transmission for Electrocardiograph [5]

By taking advantage of the ability of medium-speed and short-range wireless data transmission, this data system can be exploited for the 24-hour monitoring of portable electrocardiographs. With the recent advance of medical electronic, a electrocardiograph has been realized with 118 mm (H) × 64 mm (W) × 23 mm (D) and 150 g.

This electrocardiograph has a functionality to detect irregular pulses of the heat of the holder, as well as to memorize the waveform for 32 sec in advance of the occurrence of such irregularity. Thus once it detect any irregularity, it transfers the memorized waveform to Electrocardiograph IF, which in turn transmits it to a basestation in accordance with the protocol.

5. Concluding Remarks

This paper has devised a wireless data system composed of basestations and an IC-card type terminals. The terminal is distinctive in that a sophisticated ASIC module is incorporated with the adopted SAW-based transmitter/receiver to control the data communication as well as to lower the power consumption. This system is to be exploited for the wandering prevention of patients with senile dementia as well as for the monitoring portable electrocardiographs. Development is continuing now on the VLSI implementation of this ASIC Module, which will be reported elsewhere.

References

- [1] D. L. Ash, "SAW devices in wireless communication systems", IEEE Ultrasonics Symp. Proc., vol. 1, pp. 115-124, 1993.
- [2] D. L. Ash, "New UHF receiver architecture achieves high sensitivity and very low power consumption", RF Design, pp. 32-44, Dec. 1995.
- [3] RF Motolithics, Inc., "1995 Product Data Book", 1995.
- [4] K. Matsumura *et al.*, "A Wandering Prevention System for Patients with Senile Dementia", 17th JCMI, pp. 664-665, Nov. 1997 (in Japanese).
- [5] H. Inada *et al.*, "Development of a health information collection system for the elderly at home", 17th JCMI, pp. 678-679, Nov. 1997 (in Japanese).

```

procedure Base_station;
begin
1  Issue ENQ; /*ENQ mode*/
2  Wait until a reply is received or time-out (i.e.
   90ms of the unstable state of SAW Receiver);
   /*ID check mode*/
   if no data then back to ENQ mode;
   case receiving_data of
3    EOT: back to ENQ mode;
4    ID: if data transmission is requirement,
5       switch to data receiving mode;
6    error: issue EOT and back to ID check mode;
   end
7  Wait until a reply is received or time-out;
   /*data receiving mode*/
8  if no data then back to ENQ mode;
   case receiving_data of
9    EOT: back to ENQ mode;
10   data: issue ACK and back to data receiving
       mode;
11   error: issue NAC and back to data receiving
       mode;
   end
end

```

```

procedure Terminal();
begin /*idling mode*/
12 sleep 950ms;
13 Wake-up();
end

```

```

procedure Wake-up();
begin /*data receiving mode*/
14 Detect a pulse sequence;
15 if no pulse then back to idling mode;
16 Skip the pulse sequence until ENQ detection
   or time-out;
17 if ENQ is not detected then
   back to idling mode;
18 while ID is not issued successfully then
19   issue ID;
20   Wait until a reply is received or time-out;
21   if no data then back to idling mode;
22   case receiving_data of
23     EOT: Wait 5×N ms (N:1~16, generated by
        random number table);
24     ACK: ID transmission succeed;
25     others:
        end
   end
26 if data is not transmit then
   back to idling mode;
27 while data is not transmit successfully then
28   transmit data;
29   Wait until a reply is received or time-out;
30   if no data then issue EOT and
       back to idling mode;
31   case receiving_data of
32     ACK: data transmission succeed, issue EOT;
33     NAC: transmit data again;
34     EOT: back to idling mode;
   end
   end
end

```

Fig. 5 Protocol between Basestation and Terminal.

Analysis and Coloring of a Shotgun Cellular System

Timothy X Brown¹

Dept. of Electrical and Computer Engineering
University of Colorado, Boulder, CO 80309-0530
Tel: (303) 492-1630 Fax: (303) 492-1112
timxb@colorado.edu

Abstract

This paper analyzes cellular systems where the base stations are placed at random. Analysis compares the base station to mobile link in random systems to ideal hexagonal systems. Under log-normal shadow fading and sectorizing—but no channel planning—the random system has carrier to interference levels within 4 dB and spectral efficiency within 2 dB of the hexagonal system. A simple linear-time channel assignment algorithm reduces each of these differences by 1 dB. Thus, at a modest cost in performance, deployment costs and deployment speed can be greatly reduced.

1. Introduction

Optimal planar cellular systems place base stations (BS) in uniform size hexagonal grids. Because mountains, lakes, high rents, zoning, etc., prevent locating BS exactly at grid points, many papers have explored deviations to regular cellular grids (e.g. starting with [5]). With ever shrinking cell sizes to meet growing traffic demands and increasing resistance to BS on the part of cities and individuals, deviations in placing BS are becoming more acute. As an extreme this paper studies the limit of when the BS are placed at “random” (as in the random shot pattern of a shotgun). While easier to deploy, the performance in terms of signal quality and spectral efficiency will in general be worse than in a well-planned ideal hexagonal situation. This paper answers the question of how much worse by analyzing a cellular system with BS distributed by a purely random planar Poisson point process. It builds on results from [1], presenting a simple low complexity channel assignment algorithm, and presenting analytic and Monte Carlo results.

2. Model:

The random placement distributes BS across the plain as a 2-D Poisson point process, where the average BS density is λ and the probability of a BS in a small area dA is λdA . This will be compared with an ideal hexagonal layout. In both cases the placement is extended over the entire plain.

2.1 Model Assumptions:

Given a BS placement the following is assumed:

1. All BS have identical transmit power, antenna gains, etc., and the path loss is an inverse power law with path loss exponent ϵ .

2. Since such random cellular systems are likely in dense interference-limited systems, background and thermal noise are assumed zero.
3. Rayleigh fading (aka fast fading) is treated by micro-diversity techniques in the radio channel and is outside the scope of this paper.
4. Shadowing (aka slow fading) is modeled as independent log-normally distributed multiplicative noise, ψ , on the signal strength received from each BS [3]:

$$p(\psi) = \frac{1}{\psi \sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{\ln \psi}{\sigma}\right)^2}, \quad (1)$$

so that the signal has mean 1 and one standard deviation includes from $1/\sigma$ to σ .

5. N orthogonal channel groups are assigned to the BS according to one of several processes described below. Being orthogonal, adjacent channel interference is ignored and only co-channel interference is considered.
6. Capacity, trunking, and channel utilization issues are not addressed. As a worst case, every possible interfering channel is always active.
7. Performance is measured via the ratio of the signal BS power to the sum of all interfering BS power measured at the mobile antenna: i.e. the carrier to interference ratio (C/I). Given assumption 1, this reduces to

$$C/I = R_S^{-\epsilon} / \left(\sum_i R_i^{-\epsilon} \right), \quad (2)$$

where R_S is the distance to the signal BS and $\{R_i\}$ is the set of distances to co-channel interfer-

1. This work was completed under NSF CAREER Award NCR-9624791 and NSF Grant NCR-9725778.

ing BS. C/I performance at the BS is not considered in this paper.

8. The mobile users are uniformly distributed over the plain. They have the capability to rank the channels by C/I ratios and can choose to communicate with any BS.

2.2 Scenario Options:

This paper will analyze the C/I distribution for scenarios with the following variables and options:

- BS layout: Poisson process, or regular hexagon.
- Channel Assignment: Independent random, channel assignment algorithm, or ideal hexagonal.
- Signal Power: Path loss exponent ϵ , shadow fading standard deviation σ in dB.

The hexagon related options are used in a standard hexagonal systems. This is denoted a HEX scenario. Random layout with random channel assignment is denoted a SG (shot gun) scenario. Random layout using a channel assignment algorithm is denoted a SGA scenario. The options will be elaborated in the following sections.

3. Channel Assignment

How should N channel groups be assigned to the BS? The simplest channel assignment algorithm randomly assigns channels to a BS independent of others. In hexagonal systems, well defined regular optimal assignments are well known for many N [7]. For the shotgun BS placement no deterministic algorithm is known.

It should be emphasized that the main purpose of this paper is not to present an optimal channel assignment technique for the random cellular system. Many papers exist on the subject of channel assignment and its relationship to graph coloring [2][4][6]. Instead, a simple linear (in the number of BS) assignment algorithm that works well with this problem is presented. Algorithms using complex stochastic optimization techniques did not yield more than 1dB better results [1]. Therefore, this algorithm allows us to show the potential improvement for the random cellular system over purely random channel assignments while also having the potential to be readily applied.

The algorithm initially assigns channel groups randomly, and then pairs channel groups. For each pair it reassigns channels so that a BS and its nearest neighbor have different channels. That this is possible in linear time follows from the below theorem.

Theorem: Two colors can always color any set of points so that a point and its nearest neighbor will always have different colors. No particular metric is assumed here as long as the nearest neighbor is unique.

Proof: The nearest neighbors induce a directed (not nec-

essarily connected) graph with nodes corresponding to the points and the edges connecting from a node to its nearest neighbor. Thus all nodes have out-degree 1. Nodes are colored as follows. Choose the two nearest points in the graph. They must be connected to each other. Choose one as a root. A tree is grown as follows. By construction, connections from nodes already in the tree connect to other nodes in the tree. Thus any node adjacent to the tree must be directed toward the tree. Adding adjacent nodes one by one by induction forms a tree with all nodes having a unique directed path toward the root. When no more nodes are adjacent, the tree can be colored with two colors starting at the root with colors alternating with levels. Remove the tree's nodes from consideration and a new tree can be grown with the remaining nodes. Q.E.D.

4. Performance Results:

This section computes the C/I that at least 98% of the mobiles experience as a function of shadowing, number of channel groups, and BS placement.

4.1 Random Locations and Channel Assignments:

Analytic results are possible for random BS placement and random channel assignments. For a single channel group, Appendix A, computes the C/I distribution for all BS as:

$$\text{Prob}\left\{\frac{C}{I} > y|\epsilon\right\} = K_\epsilon y^{-2/\epsilon}, \quad (3)$$

where $y > 1$ and K_ϵ is a constant less than one that depends on the path loss exponent and plotted in Figure 3. One observation is that since ratios of distances are considered, these and all results in this paper are independent of the BS density, λ .

Assign one of N channel groups to each BS from an i.i.d. distribution with p_i the probability of choosing group i . The Poisson BS placement process implies that the group i BS can be modeled via a Poisson point process with BS density $p_i\lambda$; independent of any of the other channel groups. This implies that each channel group produces an independent C/I sample each distributed as (3). Thus, with N channels,

$$\text{Prob}\left\{\frac{C}{I} > y|\epsilon, N\right\} = 1 - \left(1 - \frac{K_\epsilon}{y^{2/\epsilon}}\right)^N. \quad (4)$$

This result is used in the next sections. For all further cases, meaningful closed form expressions were not found and Monte Carlo simulations were used. Appendix B describes the simulation method in more detail.

4.2 Shadowing:

Figure 1 plots the 98th percentile performance for 4 channels as a function of the shadow fading standard deviation. The hexagonal system deteriorates with increasing shadow fading and approaches the perfor-

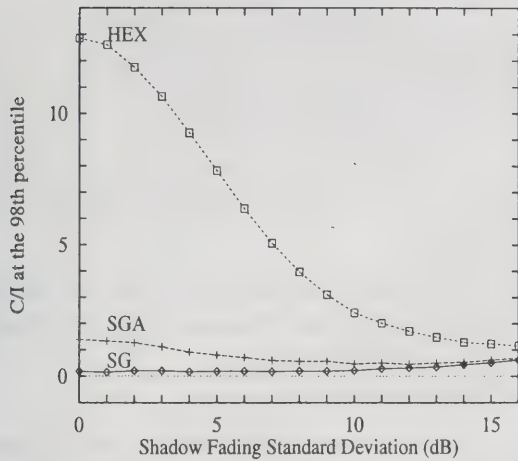


Figure 1: C/I at the 98% vs. shadow fading standard deviation with $N = 4$ channels and pathloss exponent $\epsilon = 4$ for hexagonal (HEX), random shotgun (SG) and random shotgun with channel allocation (SGA) systems.

mance of the pure random system. The random systems are insensitive to the shadow fading even when the fading component is strong.

4.3 Number of Channel Groups:

The 98th percentile performance vs. the number of channel groups with and without shadowing is shown in Figure 2. This shows that without shadowing, the hexagonal system significantly outperforms the random systems by more than 10dB. It also shows that the channel assignment algorithm improves performance by 2dB over the pure random system. With shadowing the performance is brought much closer in line. In this case channel assignment adds about 1dB to the C/I and the hexagonal system another 3dB. For a given C/I level, the required number of channels was only 2dB more.

5. Discussion:

This paper presented a cellular system design where BS are distributed randomly over the plain. This is argued to be a natural system to analyze as the limit of large perturbations to hexagonal grids.

Two random systems were analyzed with and without log-normal shadow fading. The first used random BS placement and random channel assignments. The second introduced a simple channel assignment mechanism that avoided assigning the same channel in neighboring BS.

Choosing the BS with the best C/I is a form of macro-diversity proving to be a powerful technique for avoiding poor signal quality even with no channel planning. Because of this the channel assignment algorithm improved performance by only 1–2 dB.

Without any shadow fading compared to the ideal hexagonal system the random system experiences 10dB lower C/I. On the other hand, shadow fading is an inte-

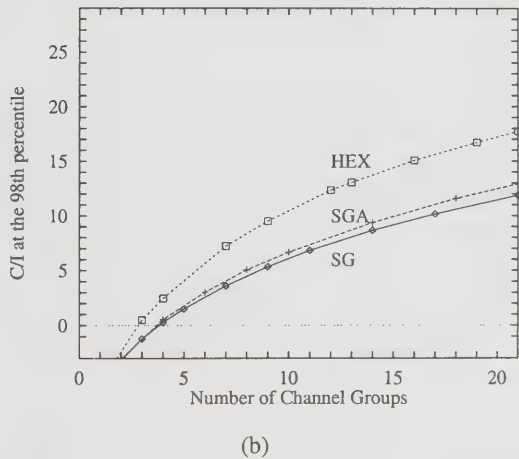
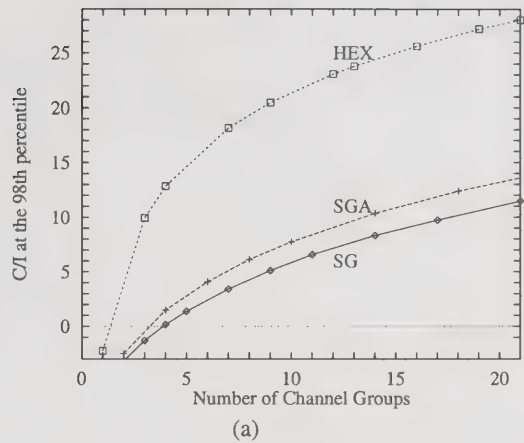


Figure 2: C/I at the 98% vs. the number of channel groups for $\epsilon = 4$. No shadowing (a), $\sigma = 10$ dB (b).

gral part of mobile communication environments. Shadow fading had a strong detrimental affect on the hexagonal system while it had little affect on the random systems. Under shadow fading C/I performance differed by only 3–4dB.

The effect of requiring more channel groups is a proportional loss in capacity per BS. This suggests that at a cost of 50% more BS, a cellular system could completely dispense with BS placement planning. Given the discrepancies from ideal already present in cellular systems, this difference may be even less. Informal talks with site acquisition specialists for cellular operators suggests that this is a price worth paying. This paper only considered the base station to mobile path. Future work would consider the mobile to base station path.

APPENDIX A: Performance of the Shotgun System

Denote the C/I distribution considering only the k nearest BS and ignoring all other BS by $\text{Prob}_k\{C/I > y|\epsilon\}$. The correct distribution is the limit of $k \rightarrow \infty$. The C/I only depends on the distance, r_i , from the user to the i th BS for all i . The probability of a random configuration

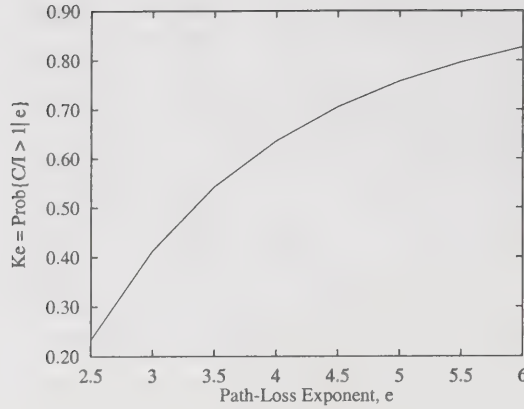


Figure 3: Factor, K_e , for computing C/I distribution as a function of ϵ in (3).

(given in [1]) is:

$$p(r_1, \dots, r_k) = e^{-\lambda \pi r_k^2} \prod_{i=1}^k 2\lambda \pi r_i \quad (5)$$

To get $\text{Prob}_k\{C/I > y|\epsilon\}$ we use (1) and integrate (5) over the set

$$\left\{ (r_1, \dots, r_k) \mid \frac{r_1^{-\epsilon}}{\sum_{i=2}^k r_i^{-\epsilon}} > y \text{ and } (r_1 \leq \dots \leq r_k) \right\} \quad (6)$$

yielding

$$\text{Prob}_k\left\{\frac{C}{I} > y|\epsilon\right\} = \int_{r_1=0}^{\infty} \int_{r_2=\max\{r_1, r_1 y^{1/\epsilon}\}}^{\infty} \dots \int_{r_k=f_k(r_1 y^{1/\epsilon}, r_2, \dots, r_{k-1})}^{\infty} e^{-\lambda \pi r_k^2} \prod_{i=1}^k 2\lambda \pi r_i dr_k \quad (7)$$

where

$$f_j(x_1, \dots, x_{j-1}) = \max\left\{x_{j-1}, \left(x_1^{-\epsilon} - \sum_{i=2}^{j-1} x_i^{-\epsilon}\right)^{-\frac{1}{\epsilon}}\right\}. \quad (8)$$

For $y \geq 1$ (note the limits for r_2), substituting $z = r_1 \sqrt{\lambda \pi y}^{1/\epsilon}$ into (7) yields two results. The first is that the distribution is independent of customer density λ . This is as expected since ratios of distances do not depend on the distance scale. The second result is that:

$$\text{Prob}_k\{C/I > y|\epsilon\} = y^{-2/\epsilon} \text{Prob}_k\{C/I > 1|\epsilon\}. \quad (9)$$

This implies that by computing $\lim_{k \rightarrow \infty} \text{Prob}_k\{C/I > 1|\epsilon\}$

once, the entire distribution can be computed for a given ϵ . Numerically integrating (9) for different k and taking limits yields Figure 3 for different ϵ .

APPENDIX B: Simulation Method:

The C/I distribution was computed by repeatedly generating cellular layouts, placing a user, and measuring the C/I of this user, each repetition representing a trial. Each

trial was generated independently of the others and repeated the following steps with parameters chosen and options from Section 2.2.

- Create a cellular layout with num_BS BS centered on the origin.
- Assign one of N channel groups to each BS.
- Place user.
- Compute the signal strength between the user and every BS using path loss exponent ϵ , and shadow fading standard deviation, σ .
- Compute the C/I for the user.

For the random cellular layout without loss of generality, the mobile user can be placed at the origin. From (5) and Bayes rule, random BS placement iteratively chooses the distance to the i th BS, r_i , using the density:

$$p(r_i | r_{i-1}) = \lambda \pi 2 r_i e^{-\lambda \pi (r_i^2 - r_{i-1}^2)} \quad (r_i \geq r_{i-1}), \quad (10)$$

where the first BS is at r_1 and $r_0 = 0$. Random channel assignment chose uniformly from the possible channel assignments. Channel assignment used the algorithm in Section 3. The other steps are straight forward.

All experiments used 1000 BS and $T = 100,000$ trials. The results are of the form $\text{Prob}\{C/I > y\}$ so that each trial is a Bernoulli trial. Since the estimate to $\text{Prob}\{C/I > y\}$ is the average of T independent Bernoulli trials denoted p , the estimate has standard deviation, $S(p, T)$:

$$S(p, T) = \sqrt{\frac{p(1-p)}{T-1}} \quad (11)$$

The error at $p = 0.98$ normalized by the maximum possible error, 0.02 is only 2.2%. Since the errors are small, error bars are not explicitly drawn on any graphs.

References:

- [1] Brown, T.X., "Analysis of Shotgun Cellular Systems," submitted to *IEEE JSAC*, June 1998.
- [2] Chamaret, B., et al., "A randomized algorithm for graph coloring applied to channel allocation in mobile telephone networks," *Proc. of the 6th Annual Conf. on Operational Research*, pp. 25–30, Oct. 1996.
- [3] Cox, D.C., "Universal Digital Portable Radio Communications," *Proc. of the IEEE*, v. 75, n. 4, pp. 436–477, Apr. 1987.
- [4] Duque-Anton, M., et al., "Channel assignment for cellular radio using simulated annealing," *IEEE T. on Vehicular Technology*, v. 42, n. 1, p. 14–21.
- [5] Macdonald, V.H. "The Cellular Concept," *The Bell System Technical Journal*, v. 58, n. 1, Jan 1979, pp. 15–41.
- [6] McEliece, R.J., Sivarajan, K.N., "Performance limits for channelized cellular telephone systems," *IEEE T. on Information Theory*, v. 40, n. 1, p. 21–34.
- [7] Rappaport, T.S., *Wireless Communications: Principles and Practice*, Prentice Hall, 1996. p. 38.

System Capacity, Latency, and Power Consumption in Multihop-routed SS-CDMA Wireless Networks

Matthew Ettus¹
Sun Microsystems
1200 Dale Avenue #17
Mountain View, CA 94040
(650) 962-8811, ettus@earthlink.net
<http://home.earthlink.net/~ettus>

Abstract

Ad-hoc wireless networking presents challenges that are different from those of tethered networks in several significant ways. In addition to high error rates and constantly varying channels, mobile communication imposes new constraints, including limited energy supplies, and the need for portability. A system for wireless networking utilizing code division multiple access (CDMA), in conjunction with spread spectrum (SS) modulation is presented. By combining SS, automatic power control and local coordination, a "collisionless," energy and spectrum efficient system can be created, which is capable of simultaneously providing high bandwidth and low latency communications. A new routing method, minimum consumed energy routing is evaluated. This new method is shown to reduce latency by 75%, reduce power consumption by 15%, and avoid congestion, in comparison with minimum transmitted energy routing. A simulator, SSNetSim, was developed to simulate the performance of these networks. By taking into account factors such as station placement, traffic patterns, routing strategies, and path loss, network performance, in terms of SNR, throughput, latency, and power consumption, is computed.

Introduction

Ad-hoc wireless networks, in which there is little or no infrastructure, present a unique challenge that is very different from those of traditional networks. Providing high throughput and low latency, efficiently using the shared spectrum, and conserving power are difficult to achieve, often conflicting goals. Scaling of these networks often leads

to poor responsiveness, collisions, and even congestion collapse.

In response to this need, a system for wireless communications using code division multiple access (CDMA), and spread spectrum (SS) modulation, is presented. Where this system differs from conventional SS-CDMA systems is in both its routing and channel access mechanisms.

¹The research detailed herein was performed while at the Institute for Complex Engineered Systems, Carnegie Mellon University. The research was supported by the Defense Advanced Research Project Agency, contracts N66001-97-C-8527 and DABT6395C0026.

Because of the use of CDMA, many stations are able to transmit at the same time, and thus no coordination across the network is necessary. The near-far problem, however, still needs to be accounted for. Through the use of local coordination only, as described in [Shep95] and [Ettus97], "collisions" may be completely avoided. By using automatic power control in conjunction with spread spectrum modulation, an energy efficient, as well as spectrum efficient, system can be created which is capable of simultaneously providing high bandwidth and low latency communications.

The primary focus of this research is to simulate the performance of large-scale spread-spectrum networks. A simulation, SSNetSim, is described, which allows the simulation of signal to noise ratios (SNR), throughput, latency, and power consumption. The simulator takes into account parameters such as number of stations, station placement, traffic patterns, routing strategies, and path loss. The results given by the simulator allow the designer to test the impact of various system parameters.

System Design

The wireless network system investigated utilizes many features which differentiate it from conventional networks. These include:

- SS-CDMA Modulation
- Cooperative Multihop Routing
- Automatic Power Control
- Local Coordination

Transmitted energy is controlled such that received power at the intended destination is held constant. Signal to noise ratio is then determined by the ratio of this received power to the power received from all other

transmissions at the same time. In an interference limited SS system, processing gain must be increased as channel SNR decreases, according to formula [1], where $SNR_{required}$ is the SNR necessary to decode the modulation used with the desired bit error rate (BER).

$$PG = SNR_{required} - SNR_{channel} \quad [1]$$

The channel SNR is not a constant, but rather follows a bell-like distribution curve. Thus, processing gain must be chosen such that a desired fraction of transmissions are received with sufficient SNR. The simulator developed is able to determine these distributions, in order to aid in the design of these systems.

In [Shep95], it was shown that in such a multihop-routed SS-CDMA network, system capacity is maximized through the use of minimum transmitted energy (MTE) routing, and that the number of stations in the network only weakly influences the overall SNR. Thus, the network may grow extremely large (millions of stations) without adversely affecting throughput.

As multihop routing is used, latency is not only determined by system bit rate, but also by how many hops are necessary for a packet to traverse the network. Unfortunately, congestion and latency rise dramatically with increasing network size (number of stations, not area), due to the increased number of stations traversed by packets en route to their destination.

A new routing method, minimum consumed energy (MCE) routing, was developed to combat these problems, as well as save energy in mobile units. Instead of routing based on minimizing the total energy transmitted along a packet's path (as in MTE),

total energy consumption is used as the metric. This includes transmitter inefficiency (which is a function of output power when automatic power control is used) and energy consumed in computer processing and encoding/decoding at each station.

Simulation

In order to demonstrate the viability of the system described previously, as well as determine how performance is affected by the various system parameters, a simulator, SSNetSim, was developed. SSNetSim allows an entire network of stations, each with its own traffic, to be accurately simulated.

As its input, SSNetSim takes information about network size, station placement, and traffic patterns. Using models for propagation and power consumption, SSNetSim is able to simulate the operation of the network as a whole. Signal to noise ratios (SNR) are computed on a packet by packet basis, as are latency, and power consumption.

Propagation modelling was essential to the evaluation of signal to noise ratios, and thus to system capacity. Various path loss models were used, including r^2 , r^3 , r^4 , and microwave oxygen-absorption band attenuation. The open design of SSNetSim allows new propagation models to be added very easily.

The routing metrics investigated include MTE and MCE, as described previously, were used. Additional methods, such as cellular (with base stations), were also looked at, but not investigated deeply. Routing in SSNetSim is very flexible, and can be adapted to match nearly any imaginable configuration.

Energy consumption in SSNetSim was modeled with the formula in equation [2],

$$E = k_1 + k_2 l + \frac{Pl}{\eta(P)} \quad [2]$$

where l is packet length, P is power for a particular transmission, and η is transmitter efficiency. Efficiency is a function of power output -- efficiency drops as power output drops below the maximum. A model of this effect for Class-B push-pull amplifiers was developed, but any desired function could be used.

Various station layouts (uniformly distribution, clustered, gaussian, etc.) were simulated, as were various traffic patterns (random destination, all-to-all, nearest neighbor, etc.). Again, these parameters could be varied within the scope of the simulator.

Results

The first test of SSNetSim was reproducing the results of [Shep95] for a baseline network with MTE routing and r^2 path loss. The two simulators produced results which matched up surprisingly well.

Under the alternative path-loss models, some interesting results were obtained. While power consumption did rise, as expected due to the need for increased transmit power in the high-loss environment, system capacity was actually improved. This may at first seem counterintuitive, but in fact it makes sense. Transmissions from distant stations (interferers) are attenuated much more, more than making up for their increased transmit power. Lower noise equates to less required processing gain, and thus more throughput for a given bandwidth.

MCE routing causes packets to take fewer hops through the network. Fewer transmissions are necessary, thus reducing congestion. Also, there are fewer heavily loaded stations, and the load is spread much more evenly, again improving the congestion problem. MCE routing has the same effect on latency.

The improvements described above do not come without some cost, however. The longer transmission distances result in more transmitted energy (which is no longer minimized), and thus slightly lower SNRs. This does decrease bit rates, but this is more than offset by the improved latency performance.

MCE routing was shown to reduce total network power consumption by about 15%, reduce latency by 75%, and reduce congestion by 75%, over standard minimum energy routing. Power consumption is also more evenly spread between stations, resulting in longer battery life for the network as a whole. Due to the increased attenuation of distant interfering stations, and a consequent rise in SNR ratios, the oxygen absorption bands actually had higher throughput than other bands.

Conclusions

When designing an ad hoc multihop CDMA network, certain design choices can help ensure optimum performance. The use of minimum consumed energy routing optimizes mobile unit battery life, as well as providing a good compromise for latency and bandwidth, to provide for high throughput.

The accurate modeling of path loss is extremely important to optimal system design. Even small changes in propagation can cause large variations in system capacity. While in

most cases propagation is out of the control of the designer, there is one exception. The choice to use the oxygen absorption bands (i.e. around 60 GHz) could greatly improve system capacity. The high loss reduces interference greatly, but is impractical for geographically sparse networks.

References

- [Ettus97] Matthew Ettus. Power Consumption and Latency as a Function of Path Loss and Routing in Multihop SS-CDMA Wireless Networks. Engineering Design Research Center, Carnegie Mellon University, December 1997. M.S. Thesis.
- [Shep95] Timothy Shepard. Decentralized Channel Management in Scalable Multihop Spread Spectrum Packet Radio Networks. Technical Report MIT/LCS/TR-670, Massachusetts Institute of Technology Laboratory for Computer Science, July 1995. Ph. D. Thesis.

A Mechanism for Matching Application Message Generation With Wireless Network Performance

Brian Rivera

Army Research Laboratory
115 O'keefe Bldg GaTech
Atlanta, GA 30332-0862
rivera@arl.mil
404 894-1810
404 894-3142(fax)

Michael Humphrey

Motorola Space and
Systems Technology Group
8201 E. McDowell Road
Scottsdale, AZ 85252

Charlene Todd

Motorola Space and
Systems Technology Group
2501 S. Price Rd
Scottsdale, AZ 85248

ABSTRACT

Because of the modular layered architecture approach to network design, the Application Layer is isolated from the characteristics of the physical network. This isolation causes the Application Layer to blindly construct packets for transmission without regard to the likelihood that these packets will be received in a timely and reliable manner at the destination node. With this type of a system, a potentially unstable condition, in which queues grow without bound, can be avoided by adding a control loop with a feedback characteristic that provides metrics on the wireless network. In any actual implementation, queue buffers are limited in size so the queues can not grow without bound.

In wireless systems, the network access delays can be rather pronounced. In such a system, knowledge of this metric can allow the Application Layer to adapt its message production rate to match the consumption rate of the network.

This paper presents a mechanism that provides feedback of channel access metric from the Data Link Layer to the Application Layer. The mechanism provides a means to exchange this information without affecting the Network and Transport Layers of the stack. This shift from the traditional layered network architecture [1,2]; while loosing some modularity and encapsulation of the stack layers, gains us more efficient use of the physical media.

INTRODUCTION*

Traditionally, network architectures are layered. This layering separates the functions which must be performed in a computer network; allowing different protocols to be used at one layer without changing the surrounding layers. The Open Systems Interconnection (OSI) protocol stack contains 7 layers, while the Internet Protocol (IP) stack contains 5 layers. The IP stack is shown in Figure 1. The datalink layer protocol we have chosen to use in this architecture study is CSMA.

The CSMA protocols are common in mobile wireless networks because they allow MAC access of a single channel without the strict synchronization and power control requirements of Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA). With CSMA, the transmitting node waits until it senses a free channel before it transmits. In high-traffic, low-data rate CSMA radios, like the military's SINCGARS system, we can expect channel access times on the order of tens of seconds for each transmission.

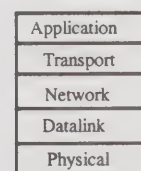


Figure 1: The IP Stack

*Prepared through collaborative participation in the Advanced Telecommunications & Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0002. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

TUNING THE MESSAGE PRODUCTION RATE

This paper studies how to tune the Application Layer message production rate to match the network's consumption rate. To understand how this concept works, assume that the Application Layer is sending a periodic message. This message may be a periodic situational awareness location update or, perhaps, a video stream or voice over IP.

The addition of video conferencing and other activities that require Quality of Service (QOS) in commercial wireless networks causes these networks to need a mechanism to control / prioritize the data that the wireless system needs. This mechanism is needed by both the transmitting and receiving station. This additional control for QOS is not within the scope of this paper.

In a wireless CSMA network, we see a queue filling up as the periodic message production rate of the Application Layer exceeds the CSMA channel access time. This condition occurs without bound if the net access delay does not improve. An "unstable" condition arises, if a majority of the nodes on the net have similar update rates and net access delays. This is because all of the queues in the net continue to grow, and exacerbate the situation.

What's worse is that the data at the front of the queue can be outdated by the data at the end of the queue. This is shown in Figure 2. In this scenario, when the Data Link layer finally gains access to the channel, old data is sent.

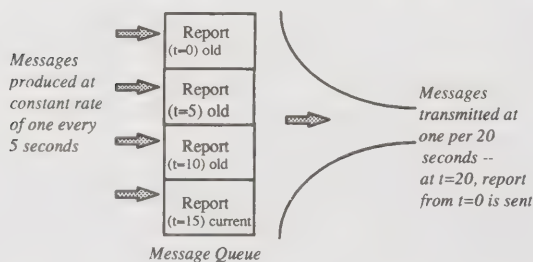


Figure 2: Message Queue Growth for Constant Five Second Inter-arrival Rate and Net Access Delay of Twenty Seconds

In a practical system, there will be other messages generated in addition to the periodic messages described here. These messages add to the overall queue growth and should be figured into the overall message production rate.

Our solution adaptively reduces the message production rate when the net access delay is high and allows the rate to increase when the delay is low. For this solution to work, we provided a communication plane framework which allows the Data Link Layer to communicate a net access delay time metric to the Application Layer. The communications plane framework is shown in Figure 3. The Application Layer insures that the queue does not overflow by producing messages at a rate equal to or slower than the net access delay time.

This communication plane can be implemented by adding MIB entries to the current MIB. These MIB entries would then be updated by the data link layer and could then be used by applications to throttle back to the current network access time.

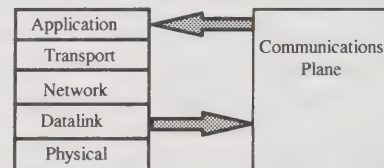


Figure 3: The Communications Plane Framework

SIMULATION RESULTS

We have performed some preliminary simulations using OPNET to validate our work. Figure 4 shows the layout of the simulated network. The mobile nodes send periodic messages the headquarters node. We gather statistics on the number of packets that arrive on time, the number that arrive late, the number of packets dropped because they expired before being transmitted, and the end to end delay of the transmitted packets.

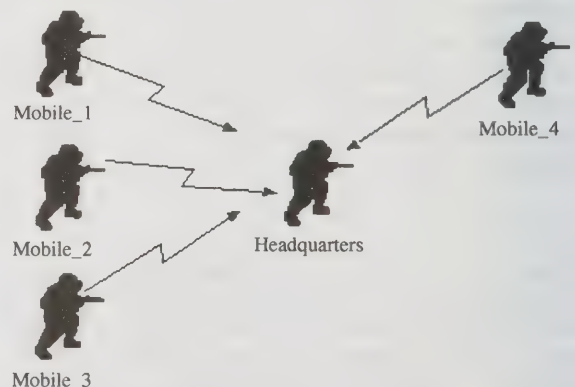


Figure 4: Network Layout

The application layer is modeled using three processes, the information generator, a packetizer, and a feedback system process. The generator represents the user creating “information” to send to the headquarters node. The packetizer addresses a datagram containing the “information” and establishes a delivery deadline and a type of “data”. In essence we are modeling information that has a finite time to live (i.e. situation awareness, video frames, voice). The feedback process throttles packet transmission to the lower layers. It limits the queuing of packets lower layer buffers by examining the status of the channel and sending packets only when the lower layers can support it. It manages queues and discards packets that have expired before transmission or that are replaced by more current “information” from the user. If more recent information arrives before the old information is sent we throw away the old information and replace it with the new. Thus the production of packets by the application layer reflects capabilities/limitations of the network.

Figure 5 shows the state machine that controls the feedback application. The **init** state initializes some variables at the start of the simulation. The **idle** state is the waiting state for the feedback process. It waits here to process packets and reacts to information from the lower layers. The **queue** state replaces old packets with newer packets and queues the different types of packets. The **send** state sends the packet with the earliest delivery deadline to the lower layers. This state also discards packets that have expired before being transmitted. This reduces the usage of scarce network resources by a packet that already is of no use to the recipient.

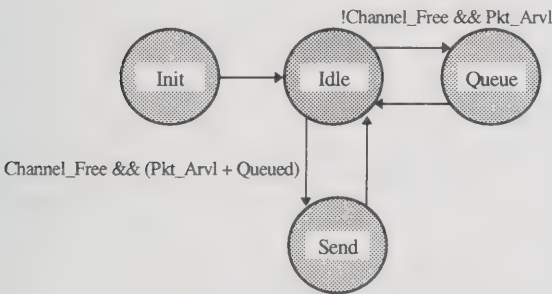


Figure 5: Process Model

For the experiment, we varied the number of mobile nodes in the network. This increased the loading on the network. For lightly loaded networks, the

feedback from the communications plane did not appreciably improve performance. This is to be expected since the messages are not being produced faster than the network can handle them. As the network load increases, the packets begin to queue up at the mobile nodes. As the packets spend more time in the queues, they are likely to expire and/or more recent data is likely to arrive. In this case, the communications plane provides significant improvement over the traditional method. As the network load increases, the communications plane continues to out perform the traditional method but still shows significant degradation due to the congestion that results from the high network load.

Tables 1& 2 show the results of one set of simulation runs. The number of packets on time is the number of packets that arrived at the headquarters node before their delivery deadline. The number of late packets are the number of packets that arrived at the headquarters node after the delivery deadline. In these simulations, the simulation duration was fixed and network load was increased. From Tables 1 & 2 it is evident that the communications plane can significantly improve network performance. For this simulated network the queues start to be significantly larger when the total packets transmitted reaches 3000. At this point the network is at capacity. As the number of generated packets increases, the network performance decreases significantly. However, the Communications Plane still provides significant improvement over the traditional system as shown by the differences in Tables 1 & 2.

Number of Packet Generated	Feedback to Application	
	Num On Time	Num Late
2000	1792	155
4000	1458	1228
6000	271	1903
8000	93	1403
10000	63	1139

Table 1: Network Performance With Communications Plane

Number of Packet Generated	No Feedback to Application	
	Num On Time	Num Late
2000	1165	142
4000	903	1616
6000	90	1555
8000	11	1085
10000	5	815

Table 2: Network Performance Without A Communications Plane

Note that the total number of on time packets and number of late packets does not equal the number of generated packets because at the end of simulation time, packets are still queued. Also, packet collisions are not shown. In the case of the communications plane, the packets that were dropped because they expired or were replaced by more recent data are not shown either.

It is also important to note that the increased performance from the communications plane depends on the relationship of packet generation to network capacity. For nodes with frequent periodic messages, the communications plane provides significant improvements, in other cases improvements are minimal.

FUTURE WORK

This simulation demonstrated the potential for the communications plane to improve network performance. There are several important issues that need to be examined including the impact of multiple hops on performance of the communications plane and effectiveness of computing transmission time and discarding packets that will expire in transit.

CONCLUSIONS

In this paper, we have shown how knowledge of network performance can help the Application Layer to make informed decisions on how to utilize the network most efficiently. Our example shows how a queue can grow without bound when periodic messages are produced at a rate faster than the net access delay time. In this condition, knowledge of the net access delay time can be used to tune the message production, matching it to the CSMA net access delay. This example also shows that simple changes to the application layer can result in dramatically better performance.

With the inter-layer feedback concepts, we can grow the concept of providing network performance information to the higher layers. Data Link Layer information, such as:

- latency
- link path loss
- link quality
- modulation waveform
- error coding
- bit-error-rate (BER)
- and Line Of Sight (LOS) obstructions

The above information can be provided to the Application Layer. We can also predict future network performance by analyzing these parameters. We can fine tune the system even more with this knowledge of expected network performance.

Information can also flow from the higher layers down to the lower layers within the communications plane. This information could be the application letting the lower layers know about QoS requirements which can be met by changing waveforms, forward error correction algorithms, or routing paths at the lower layers [3].

REFERENCES

- [1] Andrew S. Tanenbaum, "Computer Networks". Third Edition, Upper Saddle River, NJ: Prentice Hall PTR (1996).
- [2] Samuel C. Chamberlain, "Model-Based Battle Command: A Paradigm Whose Time Has Come". 1995 Symposium on C2 Research and Technology; NDU, 19 June 95.
- [3] J.E. Kleider, W.M. Campbell, "An Adaptive-Rate Anti-Jam System for Optimal Voice Communications". Proc. MILCOM '97, Monterey, CA, Nov. 1997.

*The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied of the Army Research Laboratory or the U.S. Government.

N Dimensional Orthogonal QPSK Signaling with Discrete Prolate Spheroidal Sequences

D. M. Gruenbacher and D. R. Hummels

Electrical and Computer Engineering, Kansas State University

261 Rathbone Hall, Manhattan, KS 66506

785-532-5600, 785-532-1188 (fax), grue@eece.ksu.edu, hummels@eece.ksu.edu

Abstract

A relatively new modulation method, previously referred to as Q^2 PSK, is described in further detail. An analysis for system bandwidth and bit error performance is given, along with results when the set of Discrete Prolate Spheroidal Sequences are used as the optimal pulse shaping functions. Particular attention is given to performance under channel bandlimiting, nonlinear transformations, and fading conditions.

Key words: QPSK, Discrete Prolate Spheroidal Sequences, system bandwidth, bit error performance

I. INTRODUCTION

A recent extension of QPSK called Q^2 PSK [1] doubles the number of signal dimensions from two for QPSK to four. Saha and Birdsall also mentioned that if n orthogonal pulse shapes were available then an extension of this system is possible which operates in a $2n$ dimensional signal space. Contributions of this paper include further analysis for this n dimensional QPSK signaling method, also referred to as Q^n PSK, and the use of DPSSs as the pulse shaping functions in a Q^n PSK system.

For discrete-time sequences, the set of Discrete Prolate Spheroidal Sequences (DPSSs) are known to comprise the most bandwidth-efficient set of index-limited sequences possible [2]. The DPSSs form an orthogonal set, so the set of DPSSs appears to be an excellent choice for the pulse-shaping functions of a Q^n PSK system. Although DPSSs have been difficult to generate in the past, new methods have recently been presented which make their generation easier [3]. The results given for the system bandwidth and error probabilities in the following sections have been produced with the first four DPSSs from the set defined by $N = 64$ and $W = .04$, where N is the number of DPSSs in the set and W is a bandwidth shaping factor, $0 \leq W \leq 0.5$. The power spectral densities corresponding to these sequences are shown in Figure 1.

II. SYSTEM OVERVIEW

The output of a QPSK transmitter may be represented by

$$\tilde{z}(t) = e^{j\theta^{(n)}} p(t - nT_s), \quad (1)$$

where the $e^{j\theta^{(n)}}$ represent the complex data values being transmitted and $p(t - nT_s)$ represents a pulse shaping function. For traditional QPSK the pulse

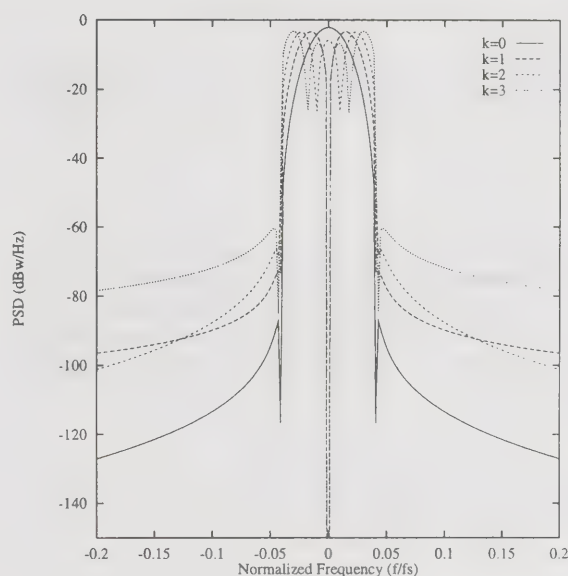


Fig. 1. Spectral densities of DPSSs $N = 64$, $W = .04$, $k = 0, 1, 2, 3$.

shaping function $p(t)$ is rectangular in shape. Non-rectangular pulse shapes are commonly used to improve the spectral characteristics of the output signal. If we have available n orthogonal pulse shapes, a Q^n PSK system can be defined by using these n pulses in n different QPSK subsystems, with the output of each summed together for the final output. For a constant transmission rate it is important to note that the symbol period increases with n as given by $T_s = n2T$. The block diagram of a Q^n PSK system transmitter is shown in Figure 2.

The biggest advantage of Q^n PSK is its potential improved spectral efficiency over similar lower dimension systems. If the n pulses being used comprise a bandwidth-efficient set of waveforms, then the bandwidth of the Q^n PSK system may decrease

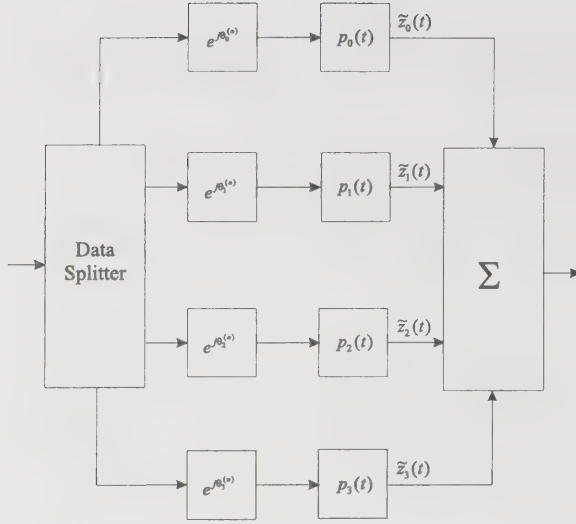


Fig. 2. Lowpass model of Q^n PSK system transmitter

appreciably as n is increased if each subsystem transmission rate is equal to $1/n$ of the original QPSK system.

III. SYSTEM POWER SPECTRA

Our objective is to find the average lowpass power spectrum of $z(t)$ indicated by $S_z(f)$, which may be found using the Wiener-Khinchin Theorem. Using methods similar to those of Titsworth and Welch [4], it can be found that the resulting average power spectrum will be

$$S_z(f) = \sum_{j=0}^{J-1} \frac{1}{T_s} P_j^*(f) P_j(f) = \sum_{j=0}^{J-1} \frac{1}{T_s} |P_j(f)|^2. \quad (2)$$

The representative power spectra of several different Q^n PSK systems were found using the results above and are shown in Figure 3. All DPSS pulse shapes used were extended in length to a final length of 512 samples.

IV. BIT-ERROR PERFORMANCE OF Q^n PSK UNDER NON-IDEAL CONDITIONS

Although it is clear the performance of Q^n PSK under the presence of additive white Gaussian noise is the same as traditional QPSK systems, the error performance of the system under other channel degradations is not so obvious. The various other channel degradations we are interested in include bandlimiting, multipath reflections, fading, and nonlinearities. We will first derive a system model under which these degradations may be considered.

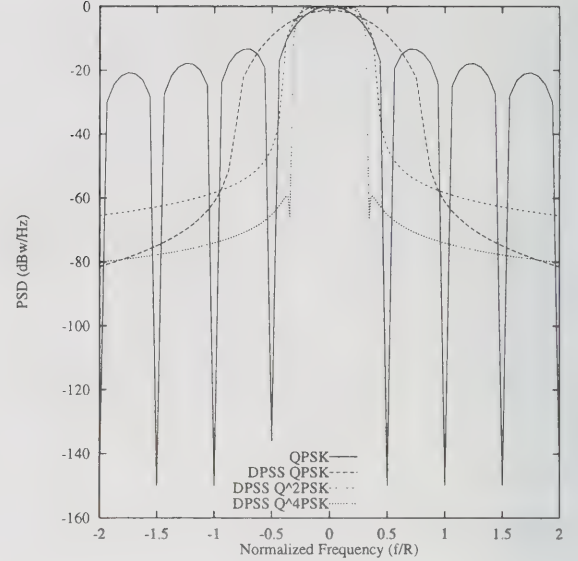


Fig. 3. Q^n PSK Power Spectral Densities

- 1) QPSK using traditional square pulse shapes,
- 2) QPSK using pulse shape DPSS $k=0$, $N=64$, $W=.025$,
- 3) Q^2 PSK with DPSSs $k=0,1$, $N=64$, $W=.025$, and
- 4) Q^4 PSK with DPSSs $k=0,1,2,3$, $N=64$, $W=.04$.

The model makes use of lowpass analysis techniques which are commonly used in studying bandpass systems [6].

The Q^n PSK transmitter in Figure 2 and corresponding matched filter receiver will be used as a guide in developing this model. Subsystems of the transmitter are referenced with j , while subsystems of the receiver are referenced with i .

Three different signals may be thought of as entering the receiver and passing through the matched filter of each subsystem. They are a result of the channel model, and are represented by:

- The direct signal, $\tilde{z}_d(t)$, with bandlimiting and nonlinearity degradations present,
- The reflected signal, $\tilde{z}_r(t)$, with the various other degradations present, and
- Channel noise, $\tilde{n}_w(t)$.

A representation for each of these signals will be found in the following sections, but first the transmitter output signal for a sequence of transmissions will be derived.

A. The Transmitted Signal

Referring to Figure 2, if the output of the j^{th} subsystem for a single transmission is denoted by $s_j(t)$, then a sequence of transmissions for that subsystem may be represented by

$$z_j(t) = \sum_{n=-\infty}^{\infty} s_j(t - nT_s). \quad (3)$$

Using the notation of complex envelopes the composite output signal may be shown to be

$$\tilde{z}(t) = \sum_{n=-N}^1 \sum_{j=0}^{J-1} A e^{j\theta_{n,j}} e^{-j2\pi f_c n T_s} p_j(t - n T_s). \quad (4)$$

Here $\theta_{n,j}$ denotes the phase of the complex data for the n^{th} transmission of the j^{th} subsystem.

B. The Direct Signal

If the signal of Equation 4 passes only through nonlinearities and channel bandlimiting before reaching the receiver, then this received signal is referred to as the direct signal $\tilde{z}_d(t)$. The appropriate matched filter response of the i^{th} subsystem of the receiver may be found to be

$$\tilde{H}_{r_i}(f) = \frac{A}{2} e^{-j\theta_i} P_i^*(f) e^{-j2\pi f_c(T_s)} e^{-j2\pi f(T_s)},$$

where θ_i is the phase of the receiver subsystem and a^* denotes the complex conjugate of a . If the response of the channel filtering is referred to as $\tilde{H}_c(f)$, then the i^{th} matched filter output at time T_s for the $n = 0$ transmission may be given by

$$\tilde{y}_i(T_s) = \frac{A^2}{2} e^{-j2\pi f_c(T_s)} \sum_{j=0}^{J-1} e^{j(\theta_{0,j} - \theta_i)} \int_{-\infty}^{\infty} \tilde{H}_c(f) P_j(f) P_i^*(f) df. \quad (5)$$

For the case where the set of pulse shaping functions $p(t)$ are doubly orthogonal and no channel filtering is present we obtain the expression

$$y_i(T_s) = \begin{cases} \frac{A^2}{2\sqrt{2}} E_p, & 0 \text{ sent} \\ \frac{-A^2}{2\sqrt{2}} E_p, & 1 \text{ sent.} \end{cases} \quad (6)$$

This is the expected output for a signal which has not been degraded in any way.

C. The Reflected Signal

The reflected signal is generated by fading and multipath of the desired or direct signal. The model for the reflected signal takes on the form

$$\tilde{z}_r(t) = R(t) e^{j\phi(t)} \tilde{z}_d(t - t_d) e^{-j2\pi f_c t_d} \quad (7)$$

where $R(t) e^{j\phi(t)}$ represents a scattering model with Rayleigh amplitude and uniform phase, and t_d is average time delay. The scattering model may be represented by

$$R(t) e^{j\phi(t)} = \sum_{k=-K}^K b_k e^{j\lambda_k t},$$

where the b_k are statistically independent zero-mean Gaussian random variables, the λ_k are constant frequency terms, and the size of K determines the accuracy of the model [5].

The mean square value of the reflected signal processed by the i^{th} section of the receiver at the sample time T_s can be found to be

$$E\{y_r^{2(i)}(T_s)\} = \sigma_r^2 = \frac{1}{4} \sum_{k=-K}^K |\tilde{y}_{r_k}^{(i)}(T_s)|^2 \frac{r_k}{(P_d/P_r)}, \quad (8)$$

where

$$\tilde{y}_{r_k}^{(i)}(t, t_d, \lambda_k, \dots) = \sum_{n=-N}^1 A \sum_{j=0}^{J-1} e^{j\theta_{n,j}} \sum_{m=-M}^M c_{m,j} \cdot \tilde{H}_i(m\omega_o + \lambda_k) \tilde{H}_c(m\omega_o) e^{j(m\omega_o + \lambda_k)t} \cdot e^{-j(n\omega_c + nm\omega_o)T_s} e^{-j(m\omega_o + \omega_c)t_d}, \quad (9)$$

the r_k are defined by a $2K + 1$ point Gauss quadrature rule, and P_d/P_r is the ratio of desired signal power to reflected signal power.

D. Channel Noise

The other source of noise is that which comes from the channel. Channel noise is usually modeled as having an additive white Gaussian distribution, with equivalent lowpass power spectrum

$$S_{\tilde{w}}(f) = 2N_o.$$

After passing through the i^{th} matched filter receiver the resulting power spectrum of the filtered noise $\tilde{n}_i(t)$ is

$$S_{\tilde{n}_i}(f) = 2N_o \frac{A^2}{4} |P_i(f)|^2. \quad (10)$$

so the variance of the channel noise at the receiver is given by

$$\sigma_n^2 = N_o E_p \frac{A^2}{4}. \quad (11)$$

E. Finding $P[e]$ using the lowpass model

For each bit permutation the conditional probability of error should be calculated for each subsystem using

$$P[e | D] = Q(y_i(T_s)/\sigma),$$

where $\sigma = \sqrt{\sigma_r^2 + \sigma_n^2}$. The overall system $P[e]$ will be the average of these subsystem performance values.

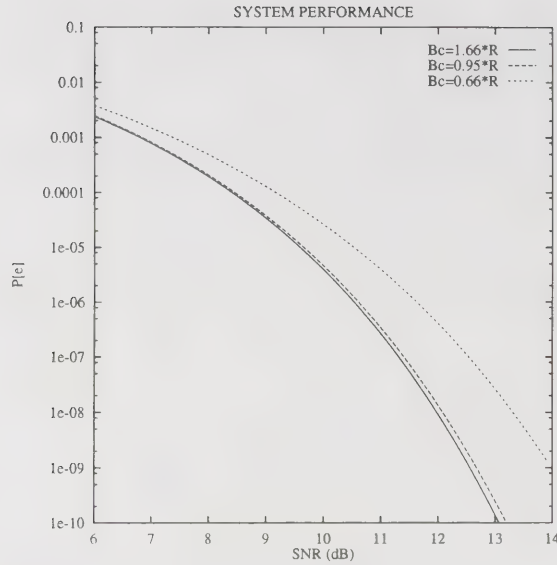


Fig. 4. Q^4 PSK Performance under Bandlimiting, $B_c/R = 0.66, 0.95, 1.66$

F. Results Using DPSSs

The results in Figure 4 show the performance of the Q^4 PSK system after bandlimiting with a four pole Butterworth filter. Figure 5 shows the system performance under nonlinear degradations with the operating level varied relative to the 1 dB compression point. Figure 6 shows the performance under fading conditions with $q P_d/P_r = 20.0$ dB, the reflected signal time delay $t_d = 1$ bit time, and fading process bandwidth B_r varied over 0.5, 1, and 2 Hz.

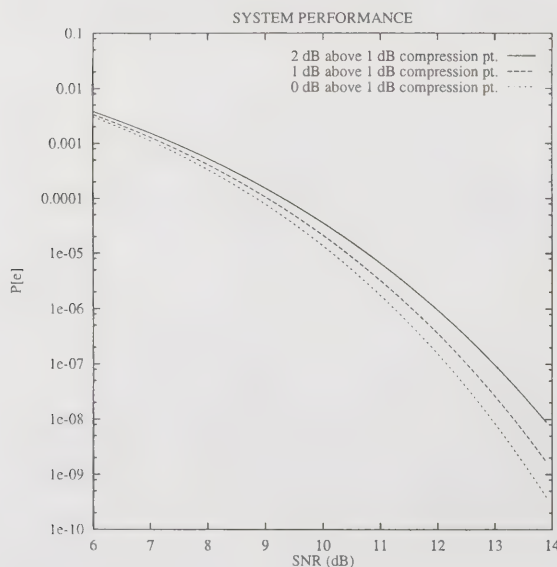


Fig. 5. Q^4 PSK Performance under Nonlinear Degradations

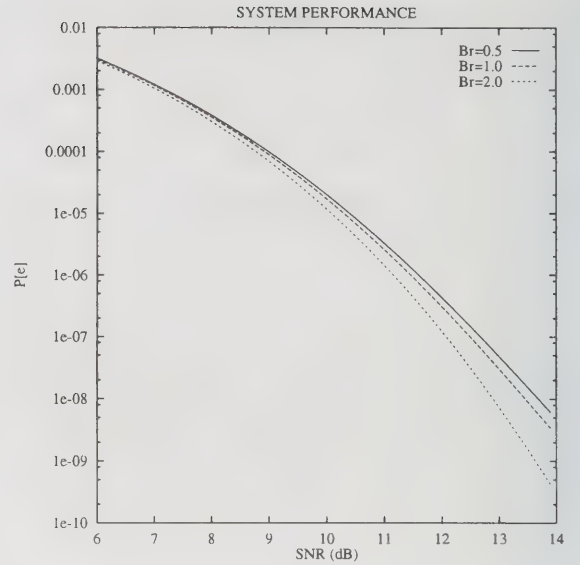


Fig. 6. Q^4 PSK Performance under Fading Bandwidth B_r

V. CONCLUSIONS

An expanded form of QPSK called Q^n PSK is shown to have potentially greater bandwidth efficiency than conventional digital communication systems, particularly when using Discrete Prolate Spheroidal Sequences (DPSSs) as pulse shaping functions. The performance of Q^4 PSK under various system degradations was studied using lowpass modeling techniques, with results showing modest improvements over current systems for most cases of channel bandlimiting.

REFERENCES

- [1] D. Saha and T. G. Birdsall, "Quadrature-Quadrature Phase-Shift Keying", *IEEE Trans. on Communications*, vol. 37, pp. 437-448, May 1989.
- [2] D. Slepian, "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty — V: The Discrete Case," *Bell System Technical Journal*, vol. 57, pp. 1317-1430, May 1978.
- [3] D.M. Gruenbacher and D.R. Hummels, "A Simple Algorithm for Generating Discrete Prolate Spheroidal Sequences", *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3277-3278, November 1994.
- [4] R.C. Tittsworth and I.R. Welch, "Power Spectra of Signals Modulated by Random and Pseudorandom Sequences", *Jet Propulsion Laboratory Technical Report*, No. 32-140, October 1961.
- [5] D.R. Hummels and F.W. Ratcliffe, "Calculation of Error Probability for MSK and QPSK Systems Operating in a Fading Multipath Environment", *IEEE Transactions on Vehicular Technology*, vol. VT-30, no. 3, August 1981.
- [6] J. Proakis, *Digital Communications*, second edition, McGraw-Hill, New York, NY, 1989.

System Performance Analysis of Impulse Radio Modulation

Fernando Ramírez-Mireles
Wireless Access Group
Glenayre Technologies

Robert A. Scholtz
Communication Sciences Institute
University of Southern California

ABSTRACT—In this paper we present a system performance analysis of impulse radio modulation using block waveform encoding signal sets. We analyze both the multiple-access performance in a channel with additive white Gaussian noise, and the single-link performance in a channel with dense multipath.

I. Introduction

Impulse radio (IR) is a spread-spectrum multiple-access (MA) technique that uses impulse signal technology to generate ultra-wideband (UWB) communication signals¹ that consist of trains of time-shifted subnanosecond impulses [1]. Data is transmitted using pulse-position-modulation (PPM) at a rate of many pulses per symbol, and multiple-access (MA) capability is achieved using spread spectrum time hopping (TH). The UWB nature of IR makes it convenient for MA communications in radio channels impaired with dense multipath (such as the wireless indoor channel [2]). With UWB TH PPM signals the dense multipath can be resolved, allowing the use of a Rake receiver [3] for signal demodulation. The radio links can be operated with reduced fade margins and the UWB TH PPM waveforms can be received by correlation detection literally at the antenna terminals, making a relatively simple and low-cost, low-power transceiver viable [1].

Although an IR system and a CDMA system operating with the same bandwidth can be shown to be quite comparable when used in a MA environment, the current impulse technology gives an advantage to IR on the basis of achievable effective processing gains for the two systems. In IR high processing gains are automatically achievable with the use of subnanosecond pulses, allowing a large number of users to be accommodated in the system.

When compared with other technologies capable of supporting G-Hz bandwidths, IR has an advantage over infrared technology since radio communication can easily penetrate the structure of buildings facilitating wireless communications. Impulse radio potentially is cheaper than millimeter wave communications for the same short-range communications environment.

The research described in this paper was done at the Communication Sciences Institute and was supported in part by the Joint Services Electronics Program under contract F49620-94-0022.

Mr. Ramírez's Ph.D. program was supported by the Conacyt Grant.

The presentation of this paper was supported by Glenayre Technologies.

The authors E-mail are {ramirez, scholtz}@milly.usc.edu.

¹The range of frequencies occupied by the UWB impulse goes from a few hundreds of Kiloherztz up to a few Gigahertz.

In this work we present a system performance analysis of IR using block waveform encoding signal sets. We analyze the multiple-access performance in a channel with additive white Gaussian noise (AWGN), and the single-link performance in a channel with dense multipath.

II. Multiple-access performance in additive noise

In this section the multiple-access performance is analyzed in terms of bit error rate for a given number of users and bit transmission rate. All the users transmit asynchronously and the desired user signal is detected using a single-channel correlation receiver for coherent detection of the symbol waveform.²

A. Channels and signals description

A.1 IR-AWGN channel model

Wireless IR channel with free space propagation conditions and AWGN. In this model the transmitted impulse is $p_{TX}(t) \triangleq \int_{-\infty}^t p(\xi) d\xi$ and the received signal is $p(t) + n(t)$.³ The noise $n(t)$ is AWGN with two-sided power density $\frac{N_0}{2}$. The signal $p(t)$ is the basic subnanosecond impulse used to convey information. It has duration T_p and energy $E_p = \int_{-\infty}^{\infty} [p(t)]^2 dt$. The normalized signal correlation function of $p(t)$ is $\gamma_p(\tau) \triangleq (1/E_p) \int_{-\infty}^{\infty} p(t)p(t-\tau) dt > -1 \forall \tau$. We define $\gamma_{\min} = \gamma_p(\tau_{\min})$ as the minimum value of $\gamma_p(\tau)$, $\tau \in (0, T_p]$.

B. PPM signals in IR-AWGN

The PPM signals received over the IR-AWGN channel consist of N_s time-shifted impulses

$$\Psi_j(t) = \sum_{k=0}^{N_s-1} p(t - kT_f - \rho_j^k), \quad j = 1, 2, \dots, M. \quad (1)$$

Here $\Psi_j(t)$ represents the j^{th} signal in an ensemble of signals, each signal completely identified by the sequence of time shifts $\{\rho_j^k; k = 0, 1, 2, \dots, N_s - 1\}$.

Each signal $\Psi_j(t)$ has duration $T_s \triangleq N_s T_f$ and energy $E_\Psi = N_s E_p$. In IR, the impulse duration satisfies $T_p \ll T_f$, where T_f is the time shift value corresponding to the frame period; and the time shift corresponding to the data modulation is $\rho_j^k \in \{\tau_1 < \tau_2 < \dots < \tau_N\}$

²This approach is analogous to the one used in [4].

³The effect of the antenna system in the UWB transmitted impulse is modeled as a derivation operation.

for some $N \geq 2$. The signals $\{\Psi_j(t)\}$ have normalized correlation values

$$\beta_{ij} \triangleq \frac{\int_{-\infty}^{\infty} \Psi_i(t) \Psi_j(t) dt}{\int_{-\infty}^{\infty} \Psi_i^2(t) dt}$$

C. TH PPM signals

In this multiple access system each user's signal is composed of a sequence of fast-hopped frame-shifted versions of one of the M possible PPM symbol waveforms $\{\Psi_j(t)\}$

$$x^{(\nu)}(t) = \sum_{m=0}^{\infty} \Psi_{d_m^{(\nu)}}(t - mN_s T_f - C_m^{(\nu)}(t)),$$

where the superscript (ν) , $(1 \leq \nu \leq N_u)$ indicates user-dependent quantities, N_u is the number of users, m indexes the transmitted symbols, $d_m^{(\nu)} \in \{1, 2, \dots, M\}$ is the m^{th} transmitted symbol, and

$$C_m^{(\nu)}(t) \triangleq \sum_{k=mN_s}^{(m+1)N_s-1} T_c c_k^{(\nu)} \phi(t - kT_f),$$

$$\phi(t) = \begin{cases} 1, & \text{if } 0 \leq t \leq T_f \\ 0, & \text{otherwise} \end{cases}$$

here T_c is a time-shift value, and $\{c_k^{(\nu)}\}$ is the pseudo-random time-hopping sequence assigned to user ν , with $0 \leq c_k^{(\nu)} \leq N_h$ for some integer N_h . To simplify the analysis, we further assume that $N_h T_c + 2(\tau_N + T_p) < T_f/2$. For a fixed T_f , the M -ary symbol rate $R_s = T_s^{-1}$ determines the number $N_s \gg 1$ of impulses that are modulated by a given symbol.

D. Multiple-access interference model

The following assumptions are made to facilitate our analytical treatment.

(a) To estimate performance without choosing a hopping code, we assume that the elements $\{c_k^{(\nu)}\}$ for $\nu = 1, 2, \dots, N_u$ and for all k , are independent, identically distributed (iid) random variables with uniform distribution on $[0, N_h]$.

(b) The transmission time differences $\tau_k - \tau_1$, $k = 2, \dots, N_u$, are iid random variables, with $\tau_k - \tau_1 \bmod T_f$ being uniformly distributed on $[0, T_f]$.

(d) We assume that the received monocycle waveform satisfies the relation $\int_{-\infty}^{\infty} p(t) dt = 0$.

E. Multiple-access system performance

When N_u links are active in this MA system, the received signal $r(t)$ can be modeled as

$$r(t) = \sum_{\nu=1}^{N_u} A^{(\nu)} x^{(\nu)}(t - \tau^{(\nu)}) + n(t) \quad (2)$$

where $A^{(\nu)}$ is the attenuation of user ν 's signal over the IR channel, $\tau^{(\nu)}$ represents clock asynchronisms between user ν transmitter and the receiver, and the signal $n(t)$ represents non MA interference modeled as

AWGN. Let's assume that the receiver wishes to determine the data modulating transmitter $\nu = 1$. The received signal $r(t)$ in (2) can be viewed as

$$r(t) = A^{(1)} \Psi_{d_m^{(1)}}(t - \tau^{(1)}) - mN_s T_f - C_m^{(\nu)}(t) + n_{\text{tot}}(t),$$

for $t \in \mathcal{T}_m$, where

$$n_{\text{tot}}(t) \triangleq \sum_{\nu=2}^{N_u} A^{(\nu)} x^{(\nu)}(t) + n(t) \quad (3)$$

includes both MA interference and thermal noise, and is assumed to be a mean-zero Gaussian random process. Standard techniques [5] can then be used to calculate the union bound on the bit error probability UBPb for coherent detection of *equally correlated* TH PPM signals. This bound is given by

$$\text{UBPb}(N_u) = \frac{M}{2} \int_{\sqrt{\text{SNRb}_{\text{MA}}(N_u)}}^{\infty} \frac{\exp(-\xi^2/2)}{\sqrt{2\pi}} d\xi,$$

where [1] [6]

$$\text{SNRb}_{\text{MA}}(N_u) = \left[[\text{SNRb}_{\text{ONE}}(1)]^{-1} + \left[\frac{1}{R_b} \frac{(2E_p[1 - \beta])^2}{2T_f \sigma_a^2 \sum_{\nu=2}^{N_u} \left(\frac{A^{(\nu)}}{A^{(1)}} \right)^2} \right]^{-1} \right]^{-1},$$

$$\sigma_a^2 = \frac{1}{T_f} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} p(t-s)[p(t) - p(t - \tau_{\min})] dt \right]^2 ds,$$

and

$$\text{SNRb}_{\text{ONE}}(1) = \frac{1}{\log_2(M)} \frac{(A^{(1)})^2 E_s [1 - \beta]}{N_o}.$$

The $\text{SNRb}_{\text{MA}}(N_u)$ and $\text{SNRb}_{\text{ONE}}(1)$ are equivalent to the output bit signal-to-noise-ratio (SNR) that one might observe in multiple-users and single-link communications, respectively. The $\beta \triangleq \frac{1+\gamma_{\min}}{2}$ is the normalized correlation value of the PPM equally correlated signals⁴. Clearly, the σ_a depend on $p(t)$ and τ_{\min} , and therefore is dependent on the signal design. Note that the bit transmission rate $R_b = \frac{\log_2(M)}{N_s T_f}$.

III. EXAMPLE

In this section we illustrate the potential MA performance of this system for a specific design. In IR modulation, the UWB received impulse $p(t)$ can be modeled by

$$p(t) = [1 - 4\pi \left[\frac{t}{t_n} \right]^2] \exp \left(-2\pi \left[\frac{t}{t_n} \right]^2 \right)$$

where the value $t_n = 0.4472$ ns was used to fit the model $p(t)$ to a measured waveform $p_n(t)$ from a particular experimental IR link. The normalized signal correlation function corresponding to $p(t)$ is

$$\gamma_p(t) = [1 - 4\pi \left[\frac{t}{t_n} \right]^2 + \frac{4\pi^2}{3} \left[\frac{t}{t_n} \right]^4] \exp \left(-\pi \left[\frac{t}{t_n} \right]^2 \right).$$

⁴For construction of PPM equally correlated signals see [7].

In this case $\tau_{\min} = 0.2419$ ns and $\gamma_{\min} = -0.6183$.

Figure (1) shows the MA performance curves. The curves represent the bit error probability with different values of M for the case in which the one-user bit SNR value $\text{SNRb}_{\text{ONE}}(1) = 10.5$ dB, so that without MA noise $\text{UBPb}(1) \simeq 4 \times 10^{-4}$ with $M=2$. The curves were calculated using $T_f = 100$ ns and $R_b = 9.6$ Kbps for each user. Perfect power control (i.e. $A^{(\nu)} = A^{(1)}$ for $\nu = 1, 2, \dots, N_u$) was assumed.

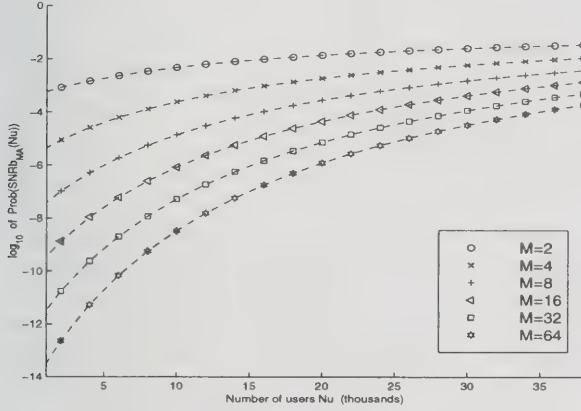


Fig. 1. The base 10 logarithm of the probability of bit error, as a function of number of simultaneous users N_u for different values of M under perfect power control conditions.

Similar curves can be calculated using high data transmission rates, in which the system is able to support hundreds of users transmitting over a Megabit per second with bit error probability in the range 10^{-5} to 10^{-8} .

IV. Single-link performance in dense multipath

In this section we make an assessment of the performance of IR modulation in an indoor multipath environment with detection using a Rake receiver. In the present analysis we assume one user and perfect synchronization. Under these circumstances, the spread spectrum time-hopping sequence modulation has no effect on the correlation properties of the communication signals, and will be omitted in the expressions defining the signals and their correlation values.

A. IR-MP channel description

Wireless indoor IR multipath channel. In this model the transmitted signal is $p_{\text{TX}}(t)$ and the received signal is $\sqrt{E_p} \tilde{p}(u, t) + n(t)$. The pulse $\sqrt{E_p} \tilde{p}(u, t)$ is a time spreaded version of $p(t)$, with duration $T_m \gg T_p$ and average energy \bar{E}_p . The u indexes an event taking place in the sample space of a certain random experiment. Hence, $\tilde{p}(u, t)$ is a random process.⁵ For performance analysis purpose, it will be assumed that the IR-MP

⁵For the IR-MP channel the random experiment is a measurement experiment performed in an office building where $\tilde{p}(u_o, t)$ denotes the IR-MP channel pulse response to $p_{\text{TX}}(t)$, measured at position u_o in the absence of noise.

channel can be characterized by the ensemble of pulses responses

$$\{\tilde{p}(u_o, t)\}, u_o = 1, 2, \dots, u_*.$$

The normalized signal correlation function of $\tilde{p}(u_o, t)$ is

$$\gamma_{\text{MP}}(u_o, \tau) \triangleq \frac{\int_{-\infty}^{\infty} \tilde{p}(u_o, t) \tilde{p}(u_o, t - \tau) dt}{\int_{-\infty}^{\infty} [\tilde{p}(u_o, t)]^2 dt}$$

B. PPM signals in IR-MP

The PPM signals received over the IR-MP channel consist of N_s time-shifted impulses

$$\tilde{\Psi}_j(u_o, t) = \sum_{k=0}^{N_s-1} \tilde{p}(u_o, t - kT_f - \rho_j^k), \quad j = 1, 2, \dots, M.$$

Each signal $\tilde{\Psi}_j(u_o, t)$ has duration $\tilde{T}_s \triangleq N_s T_m$, energy $\tilde{E}_{\Psi}(u) = \int_{-\infty}^{\infty} [\tilde{\Psi}(u_o, \xi)]^2 d\xi$, and average energy $\bar{E}_{\Psi} = N_s \bar{E}_p$. For the multipath analysis we further assume the channel varies slowly with respect to \tilde{T}_s , and that $\tau_N + T_m < T_f$. The signals $\{\tilde{\Psi}(u_o, t)\}$ have normalized correlation values

$$\tilde{\beta}_{ij}(u_o) \triangleq \frac{\int_{-\infty}^{\infty} \tilde{\Psi}_i(u_o, \xi) \tilde{\Psi}_j(u_o, \xi) d\xi}{\int_{-\infty}^{\infty} [\tilde{\Psi}_i(u_o, \xi)]^2 d\xi}.$$

C. Single-link multipath performance

When the signal $\int_{-\infty}^t \Psi_j(\xi) d\xi$ is transmitted over the IR-MP channel, the received signal in a symbol interval $\tilde{T}_s = N_s T_f$ can be written

$$r(u_o, t) = \tilde{\Psi}_j(u_o, t) + n(t), \quad 0 \leq t \leq \tilde{T}_s$$

Conditioned on the measurement location u_o , the received signal consists of the communication signal $\tilde{\Psi}_j(u_o, t)$ plus AWGN (of course, for different u_o , the sets of received signals are different). For every u_o , the optimum receiver is a bank of filters matched to the M signals $\tilde{\Psi}_j(u_o, t)$, $j = 1, 2, \dots, M$. The problem is that this receiver must be able to match the random variations in the received signal for every possible value u_o . With this motivation, we introduce the concept of *perfect Rake* (PRake) receiver, a super Rake receiver that has an unlimited number of correlation resources and is able to construct a reference signal $\tilde{\Psi}_j(u_o, \tilde{T}_s - t)$ that is perfectly matched to the signal received $\tilde{\Psi}_j(u_o, t)$ over the multipath channel. For every u_o , performance analysis using the PRake receiver can be calculated, and the average performance can be obtained by averaging over all values of u_o . The PRake receiver provides a theoretical bound for the best performance attainable.

Conditioned on $u = u_o$, the union bound on the bit error probability Pe for *equally correlated* UWB PPM signals detected with the PRake receiver can be written

$$\text{UBPb}(u_o) = \frac{M}{2} \int_{\sqrt{\text{SNRb}_{\text{MP}}(u_o)}}^{\infty} \frac{\exp(-\xi^2/2)}{\sqrt{2\pi}} d\xi,$$

where [8]

$$\text{SNRb}_{\text{MP}}(u_o) = \frac{\tilde{E}_{\Psi}(u_o)}{N_o} \log_2(M)(1 - \tilde{\beta}(u_o)).$$

The $\tilde{E}_{\Psi}(u_o)$ value accounts for variations in the received signal energy due to fading caused by multipath. The $\tilde{\beta}(u_o) \triangleq \frac{1 + \gamma_{\text{MP}}(u_o, \tilde{\tau}_{\min})}{2}$ is the normalized correlation value of the UWB PPM equally correlated signals. The $\tilde{\beta}(u_o)$ value accounts for distortions in the signal correlation function caused by multipath [9].

The value $\text{UBPb}(u_o)$ is conditioned on the event $u = u_o$, and depends on $\tilde{\Psi}_j(u_o, t)$ (i.e., $\tilde{p}(u_o, t)$), the channel waveform measured in the absence of noise at position u_o . Taking the expected value $\mathbf{E}_u\{\cdot\}$ with respect to u over all positions

$$\overline{\text{UBPb}}\left(\frac{\tilde{E}_{\Psi}}{N_o}\right) = \mathbf{E}_u\{\text{UBPb}(u)\}$$

where

$$\left(\frac{\tilde{E}_{\Psi}}{N_o}\right) \triangleq \mathbf{E}_u\{\text{SNRb}_{\text{MP}}(u)\}$$

is the average received symbol SNR.

D. EXAMPLE

The expected value $\overline{\text{UBPb}}\left(\frac{\tilde{E}_{\Psi}}{N_o}\right)$ can be approximated by the sample mean value

$$\overline{\text{UBPb}}\left(\frac{\tilde{E}_{\Psi}}{N_o}\right) \approx \frac{1}{u_*} \sum_{u_o=1}^{u_*} \text{UBPb}(u_o)$$

Figure (2) shows the curves for $\overline{\text{UBPb}}\left(\frac{\tilde{E}_{\Psi}}{N_o}\right)$, $M = 2, 4, 8, 16, 32, 64$, using the same impulse shape $p(t)$ used in the MA example. In this case the value $t_n = 0.7531$ ns was used to fit the model $p(t)$ to the measured waveform $p_m(t)$. The UWB impulse $p_m(t)$ is a unitary-energy template with duration $\tilde{T}_p = 1.5$ ns that was taken from a multipath-free and noise-free measurement. The resultant correlation function of $p(t)$ has a minimum $\gamma_{\min} = -0.6183$ as before, but the minimum is attained at the time shift value $\tilde{\tau}_{\min} = 0.7531$ ns. The channel pulse responses $\tilde{p}(u_o, t)$, $u_o = 1, 2, \dots, u_* = 392$ come from signal propagation data recorded in an UWB measurements experiment [10].

V. Conclusions

Results from this analysis show that for applications requiring high data rate (1024 Kbps) combined with low probability of bit error (10^{-6}), IR modulation is potentially able to support hundreds of users. Similarly, for applications requiring low data rate (9.6 Kbps) and moderate probability of error (10^{-4}), IR is potentially able to support dozens of thousands of users. In either case, the combined transmission rates give a transmission capacity near to 500 Megabits per second using receivers of moderate complexity.

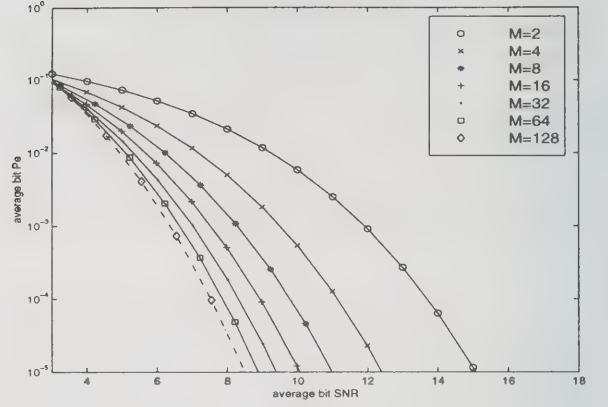


Fig. 2. The $\overline{\text{UBPb}}\left(\frac{\tilde{E}_{\Psi}}{N_o}\right)$ Curves for $M = 2, 4, 8, 16, 32, 64$ signals.

The real payoff of IR will be in wireless transmission in multipath channels. This analysis show that for a single-link case and symbol error probability of 10^{-3} , the performance in the presence of multipath using an idealized Rake receiver is, on average, a few dB worse than performance in the absence of multipath.

These results support the conclusion that IR will have good multiple access performance even in dense multipath channels.

VI. Acknowledgments

The UWB signal propagation data used in the numerical example was kindly provided by Dr. Moe Z. Win from measurements made at Time Domain Corp.

References

- [1] R. A. Scholtz, "Multiple Access with Time Hopping Impulse Modulation," invited paper, Proceedings of Milcom'93, Dec. 1993.
- [2] H. Hashemi, "The Indoor Radio Propagation Channel," Proceedings IEEE Vol. 81, No. 7, July 1993.
- [3] R. Price and P. E. Green Jr., "A Communication Technique for Multipath Channels," Proc. IRE, March 1958, pp. 555-570.
- [4] C. L. Weber, G. K. Huth and B. H. Batson, "Performance Considerations of Code Division Multiple-Access Systems," IEEE Trans. on Vehicular Technology, Vol. VT-30, No. 1, pages 3-9, February 1981.
- [5] R. M. Gagliardi, *Introduction to Telecommunications Engineering*, John Wiley and Sons, 1988.
- [6] F. Ramírez-Mireles and R. A. Scholtz, "Multiple-Access with Time Hopping and Block Waveform PPM Modulation," in Proceedings IEEE ICC'98, June 1998.
- [7] F. Ramírez-Mireles and R. A. Scholtz, "Time-Shift-Keyed Equicorrelated Signal Sets for Impulse Radio M-ary Modulation," in Proceedings WIRELES'98, July 1998.
- [8] F. Ramírez-Mireles and R. A. Scholtz, "Performance of Equicorrelated Ultra-Wideband Pulse-Position-Modulated Signals in the Indoor Wireless Impulse Radio Channel," in Proceedings IEEE PACRIM'97, Aug. 1997.
- [9] F. Ramírez-Mireles, M. Z. Win and R. A. Scholtz, "Signal Selection for the Indoor Impulse Radio Channel," in Proceedings IEEE VTC'97, May 1997.
- [10] M. Z. Win and R. A. Scholtz, "Ultra-Wide Bandwidth (UWB) Signal Propagation for Indoor Wireless Communications," in Proceedings IEEE ICC'97, June 1997.

In-Band Intermodulation Measurement and Simulation of a Local Multipoint Distribution System

Thomas Wong

**Department of Electrical and Computer Engineering
Illinois Institute of Technology
Chicago, Illinois 60616
U. S. A.**

Tel: (312) 567-5796 Fax: (312) 567-8976

Email: twong@ece.iit.edu

Abstract

Measurement results of the in-band intermodulation performance of a millimeter-wave communication system developed for Local Multipoint Distribution applications are reported. The system employed traveling-wave tubes as the headend power devices for asymmetrical information transfer. Implications of the presented results on system design, and procedures to incorporate the data for system simulation tools to predict baseband characteristics in terms of representations such as constellation plots are discussed.

I. Introduction

Among the characteristics of a broadband wireless communication system, in-band intermodulation is the one that has direct impact on the link budget, total information throughput, strategy for frequency reuse, and consequently the total system cost. In this paper, we report on measurement results of the in-band intermodulation (IM) performance of a millimeter-wave communication system developed for Local Multipoint Distribution System (LMDS) applications [1]. When asymmetrical rates of information transfer are required for a multipoint distribution system, power devices with high-saturated power output such as the traveling-wave tube (TWT) offer considerable advantages for use at the headend in the aspects of cost, system integration and maintenance. While vacuum devices usually possess soft limiting characteristics and are more

tolerant to being over-driven, they also exhibit nonlinear behavior that will affect the performance of the communication system. For a multicarrier transmission system typical of the schemes employed in LMDS, the intermodulation performance can be directly measured with the system excited by a multicarrier IF and the IM level noted by observing the spectrum at the receiver. Extensive investigations on the IM characteristics of the TWT and their effects on multicarrier satellite communication systems have been carried out previously [2]. Recent advances in nonlinear system simulation tools allow the incorporation of such IM measurement results with the tools to predict the performance of a communication link at baseband for the received signal [3]. Such information is useful to the system designer in answering questions such as whether a certain modulation and coding scheme will yield acceptable performance for a

given combination of transmitter power, antenna gains, and receiver sensitivity.

II. System Description

A schematic diagram of the system employed in this study is shown in Figure 1. To replicate the propagation loss in an open-space link, considerable attenuation in conjunction with a 20 dB coupler were placed at the transmitter output. In a typical receiver designed for LMDS, the IF is often chosen in the 950 - 2050 MHz range. Accordingly, an LO of 26.5 GHz was employed. The spectrum of the down-converted signal as displayed by the spectrum analyzer at nominal system operation conditions is shown in Figure 2. A fundamental issue for system characterization with multitone excitation is the variation in in-band intermodulation level as the system is gradually driven into saturation. The IM level at the receiver IF for different levels of transmitter output is shown in Figure 3. To enable the construction of behavioral models to characterize the nonlinear properties of the system, a network analyzer was employed, as shown in Figure 4. To meet with the requirement for S-parameter measurements, the LO frequency of the receiver was selected to be identical to that of the transmitter. The most important property revealing the nonlinear characteristics of the communication link in the context of such measurements is the forward transfer factor, S_{21} . Frequency sweeps of the system forward transfer factor (IF to IF) for two transmitter-power settings are shown in Figure 5. All data presented have been repeated over a period of 6 months, during which the system has been put in regular operation under nominal conditions.

III. Discussions

It can be seen from the output spectrum of the receiver that the system in this study can support the distribution of a variety of information contents at the level of 25 carriers per TWT with substantial link margin in weather conditions typical of the Chicago metropolitan area, employing various modulation schemes such as FM, FSK, BPSK, and QPSK. The use of higher spectral efficiency modulation schemes can be implemented when scaling measures such as sectoring and spectrum allocation are exercised accordingly. The use of robust error correction codes further reduces the effective interference level due to system nonlinearity.

Much attention has been given to the measurement, modeling and compensation for the nonlinear characteristics of TWT in satellite communications. The general formulation [2] developed in the early days imposed no restriction on the number of carriers and total bandwidth of the device. From a practical standpoint, the modeling of the nonlinearity in the device at the time when closed form expressions were preferred due to limitations in computation technology imposed considerable constraints on the complexity and hence the scope of the modeling effort. With the development in modern system representations such as the Volterra-series, generalized power series and continuous intermodulation analysis method, more refined simulation tools can be developed. In this context, two aspects of nonlinearity exhibited by a broadband communication system with instantaneous bandwidth found in LMDS warrant special attention. The first one can be revealed by observing that the two curves shown in Figure 5 do not have the

same shape. There is corresponding shift in the phase of the forward transfer factor (not shown.) This is a manifestation of the nonlinear behavior of the power device that often has been described in terms of AM-AM and AM-PM conversions, coupled with the effects of other quasi-linear devices such as mixers (as far as in-band transfer properties are concerned) and the frequency dependence of linear components, such as band-pass filters, in the communication link. The second one can be observed from the competition for power between the carriers as a selected group is given more strength. This may not be an important issue for constant envelope systems, but must be addressed in amplitude-dependent systems.

In summary, in-band intermodulation measurements on an LMDS developed for asymmetrical information transfer were presented.

Their implications on system design and interface with recently developed simulation tools for prediction of system performance were discussed.

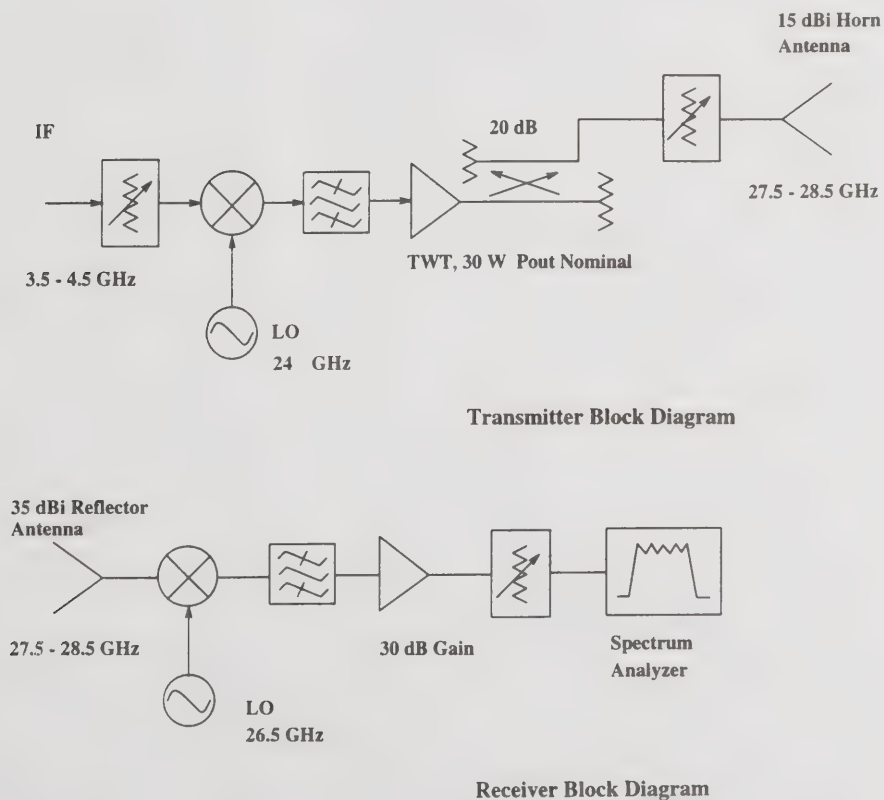
IV. Acknowledgment

The contribution of Fang Liu at the early stage of this project is gratefully acknowledged.

V. References

1. T. Wong, "An equipment developer's perspective of LMDS technology," *ICM Conference on Wireless Cable and LMDS*, August 1997, Chicago
2. O. Shimbo, "Effects of intermodulation, AM-PM conversion, and additive noise in multicarrier TWT systems," *Proc. IEEE*, vol. 59, pp. 230-238, February, 1971.
3. See for e.g., "Serenade system simulator," *Ansoft Corporation*, 1998

Figure 1. Equipment arrangement for IM measurement.



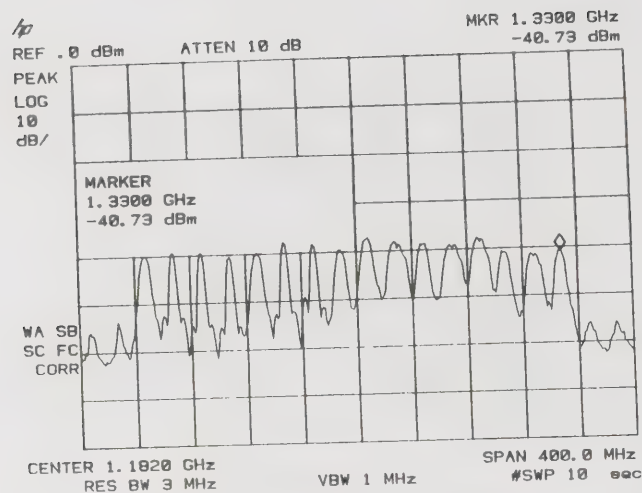


Figure 2. Receiver IF spectrum for 16 consecutive channels at nominal transmitter output of 30W.

Figure 3. In-band IM level for different transmitter output with 16 consecutive carriers at 20 MHz intervals.

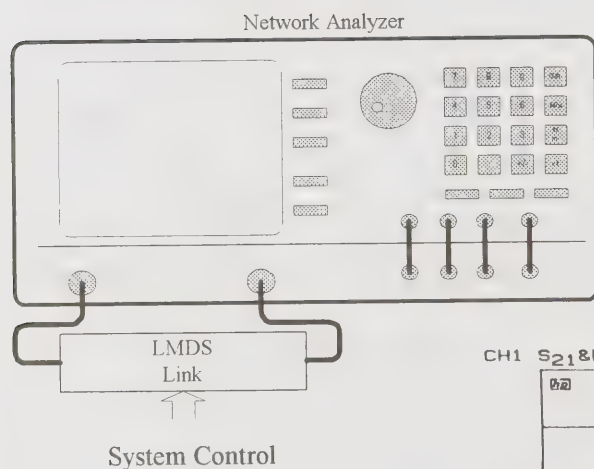
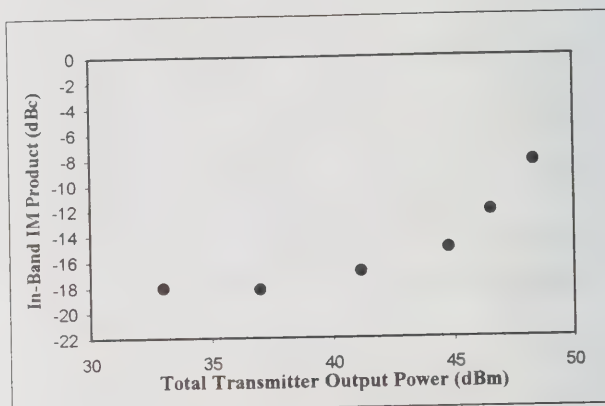
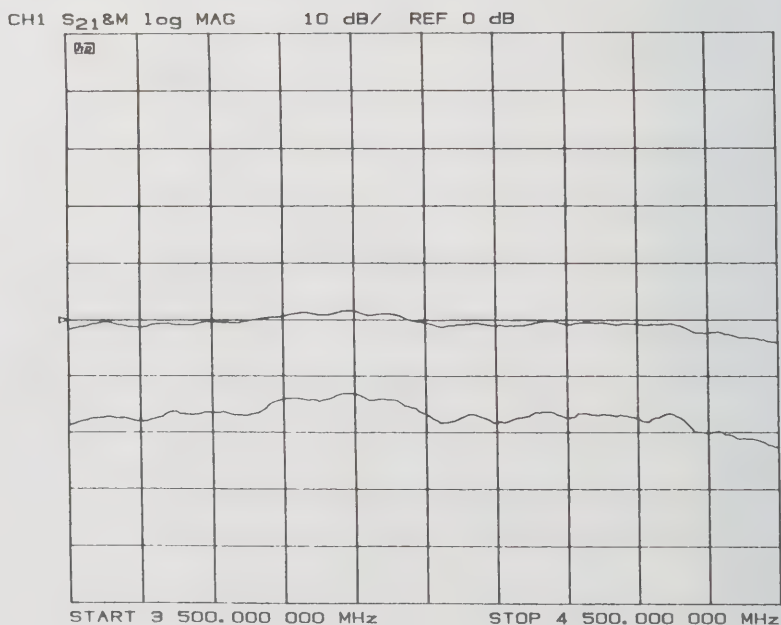


Figure 4. System forward transfer factor measurement employing a network analyzer.

Figure 5. Forward transfer factor, $|S_{21}|$, of the system with transmitter output of 30W and 0.8W at midband.



Noise in Self-timed and Synchronous Implementations of a DSP

W.A. Lien, P. Day, C. Farnsworth, R. Glibbery, D.L. Jackson, J. Liu, N.C. Paver, J. Chappel, K. Kozma¹

Cogency Technology Incorporated

144 Front Street West, Suite 580,

Toronto, Ontario, Canada M5J 2L7

Phone: (416) 217-0250, Fax: (416) 217-0256, Email: lien@cogency.com

Abstract

This paper proposes the use of self-timed design as an elegant and cost effective solution to control noise. The importance of reducing the high frequency components of the switched current is demonstrated from a review of noise sources in electronic systems. The paper includes a comparison between synchronous and self-timed systems based on the implementation of a DSP and clearly demonstrates the benefits of self-timed design.

1. Introduction

Latest designs for the embedded controller mobile market use complex logic for functions such as base-band RF, internet access, graphics display and general purpose computing. The number of transistors and operating frequency of such circuits has kept pace with advances in process technology, leading to an increase in switching noise and a corresponding increase in noise related problems [1].

Recent academic research has focussed on areas such as modeling noise in substrates [2], reduction of noise from simultaneously switching I/O [3], and design tools for noise optimal power routing [4] or cell placement [5]. However, the very nature of synchronous systems limits the reduction in noise that may be achieved via these techniques. For example, noise is commonly lowered by reducing the number of logic gates that switch simultaneously (separation in time) or by physically separating sensitive circuits (separation in space). Synchronous systems offer little latitude in reducing noise by separation in time as the clocked circuits, combined with aggressive minimum clock skew requirements, inherently guarantee the simultaneous switching of logic throughout a VLSI layout. In addition, the noise transport mechanism for low resistivity substrates common in CMOS processes, is primarily through indirect paths such as the low impedance power routing, and not by direct coupling through the substrate [3]. Therefore, physical separation of noise sources from noise sensitive circuits or barrier methods, such as guard rings, are less effective options for isolating switching noise of clocked logic in CMOS processes.

Self-timed logic offers a viable alternative for the reduction of noise. Unlike synchronous design, self-timed systems propagate data using handshake signal-

ling between a sender and a receiver, eliminating the need for a global clock. Switching of logic is then governed by the processing delay between communicating components rather than a central timekeeper. Intuitively, one would expect the number of simultaneously switching outputs in self-timed systems, and therefore the peak switched current, to be less than an equivalent synchronous implementation. Furthermore, the fact that self-timed circuits only consume energy when doing useful work implies lower power consumption and less noise [6][7].

2. Self Timed Design - An Overview

Self-timed design involves the use of functional units with a simple handshake interface for the sending or receiving of data and control status. Combining such functional units allows the construction of complex self-timed systems. Figure 1 shows two functional units; one unit is defined as a sender who wishes to pass data; the other is the receiver who wishes to receive data.

Data to be communicated between the sender and receiver is bundled with two control wires, request, Req, and acknowledge, Ack. This interface can have many differing timing protocols associated with the data communication [8][9]. Figure 2 shows one possible communication protocol, the four phase "early" protocol. Here, the sender prepares the data and then notifies the receiver that data is available by asserting Req. The

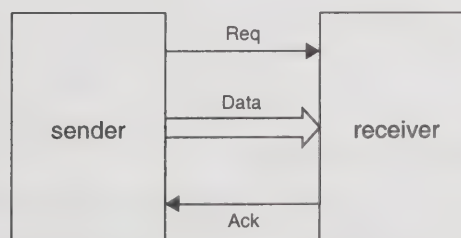


Figure 1. Self-timed bundled data interface

1. Xentec Inc., 300-2908 South Sheridan Way, Oakville, Ontario, Canada, L6J 7J8, (905) 829-8889

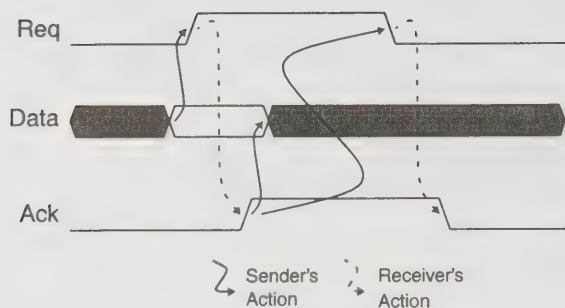


Figure 2. 4-phase "early" bundled data handshake protocol

receiver consumes the data and asserts the acknowledge signal, Ack, allowing the sender to remove the data and begin preparing the next data packet. To complete the transfer a recovery phase is then initiated by the sender by de-asserting Req. The communication cycle completes when Ack returns to zero, after which a new transfer can begin. Communication between sender and receiver must follow this strict ordering of events. There is however, no upper bound on the delays between consecutive events.

The self-timed DSP design described in this paper uses a 4-phase bundled data interface protocol for communication between blocks such as a register file, memory, multiplier, and ALU.

3. Self-Timed DSP

Cogency Technology has designed a commercial implementation of a self-timed DSP [6]. The self-timed design is instruction set compatible with a synchronous version developed by LG Semicon [10]. The DSP is a 16-bit fixed point architecture designed for Fax/Modem applications. Both devices were implemented in the same 0.6 μm technology, have almost identical die sizes, and operate at 5V.

Results from a transistor level simulation using EPIC Powermill were used to derive noise spectrums for both devices running code to implement a Hilbert filter, at an equivalent frequency of 30 MHz. The code executes numerous multiply accumulate operations as well as the appropriate looping, and is therefore, a realistic representation of a typical DSP work load. Testing of the manufactured devices confirmed average current measurements were within 5% of the Powermill simulation.

4. Noise Comparisons

Using the synchronous and self-timed implementations of the DSP, noise comparisons can be made for substrate coupling, packaging noise, and electromagnetic interference (EMI).

Substrate Coupling

Substrate noise is an example of conductive coupling that results from the nonzero dielectric constant and conductivity of the substrate leading to current flow between a switching transistor (noise source) and another device. Mixed signal and low voltage circuits are very susceptible to substrate noise problems.

A detailed discussion of substrate coupling is beyond the scope of this paper and can be found in [2][11]. This paper focuses on the low resistivity substrate of a typical CMOS process. In such a process, noise is injected directly into the power routing via the substrate inductance, rendering the use of guard rings and physical separation ineffective.

The transfer function between the injected current and the received voltage, formed by the substrate coupling and package inductance, is a complex high pass filter. This high pass substrate coupling is responsible for the problem noted by Texas Instruments in the design of an A/D converter [3]. High frequency injected current caused under-damped oscillations across the package inductance resulting in extended output ringing.

Substrate noise can be avoided by modifying the transfer function or by reducing the injected noise. Modification of the transfer function requires the use of expensive process and packaging technologies not viable for mobile products designed for the consumer marketplace. Therefore, reduction of noise at the source, particularly of the high frequency spectral components of the injected current, is critical.

In a synchronous design, the injected current has a large high frequency component caused by the simultaneous switching of components on the clock edge, combined with the high frequency components of the global clock signal itself. In a self-timed design only the currently active components are switching, and switching occurs in response to local handshake signals rather than a global clock. This lack of coherent switching results in an injected current that is both lower in magnitude and lower in high frequency composition.

The injected current is a percentage of the total current. By assuming that this percentage is the same for synchronous and self-timed designs, the substrate noise in the two designs can be compared by multiplying the total current by the substrate noise transfer function.

The substrate noise spectrums for the synchronous and self-timed DSP's are shown in Figure 3. The synchronous device exhibits significant noise at frequencies above 100 MHz. This is expected as the substrate has a high pass transfer function and therefore passes the high frequency components present in synchronous systems. Conversely, substrate noise for the self-timed DSP is

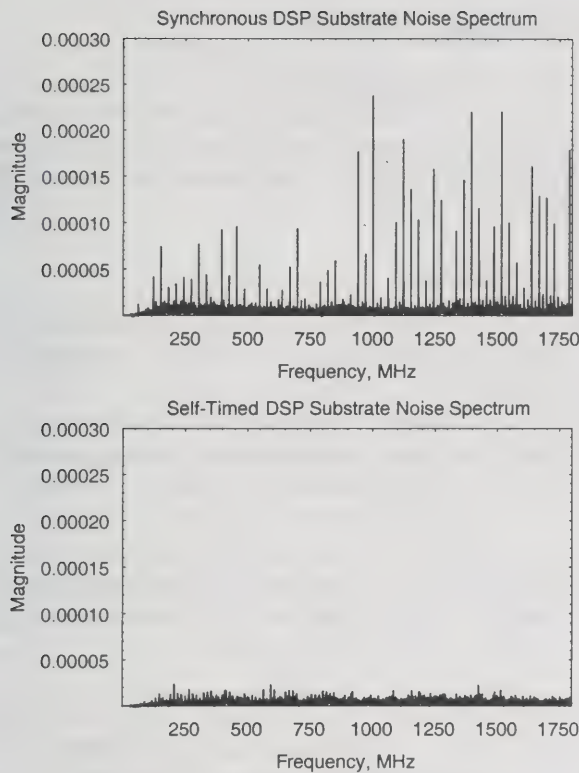


Figure 3. DSP Substrate Noise Spectrum

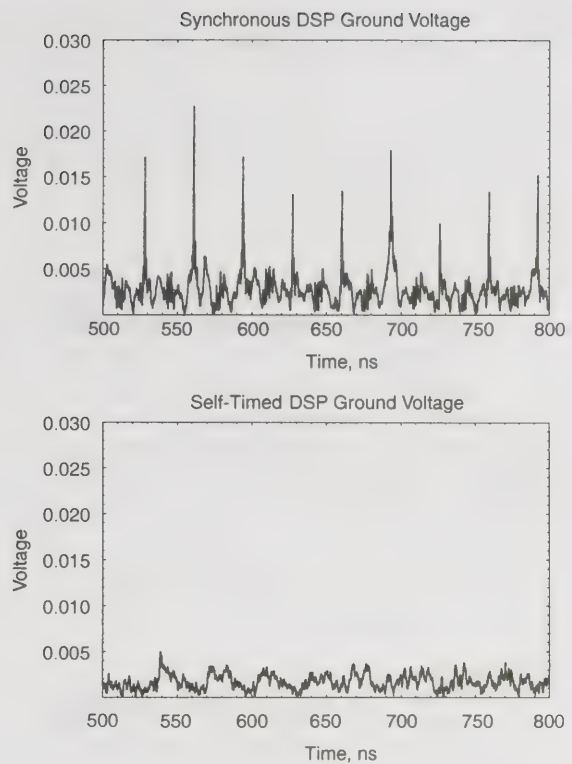


Figure 4. Ground Voltage

essentially broadband white noise due to the absence of coherent high frequency switching. It is also notably lower in magnitude.

Packaging Noise

The package inductance figures prominently in the behavior of substrate noise, as noted above. The $L di/dt$ voltage drop across the package also generates common impedance noise (ground bounce) which in the past has been countered through the use of expensive packaging technologies to reduce the inductance. Alternatively, packaging noise can be reduced by minimizing the high frequency spectral components of the current through the use of self-timed design techniques.

Figure 4 shows that the lack of high frequency current components for the self-timed DSP results in the elimination of large transients in the ground voltage when compared to the synchronous version.

Electromagnetic Interference

EMI is the most important form of noise in VLSI circuits for designers of wireless systems hoping to isolate digital noise from sensitive analog RF circuitry. Radiation from digital circuits occurs as a result of current flowing around loops formed by the conductors of the circuit. These loops act as antennas, radiating an electromagnetic field.

The far field of a semiconductor device can be approximated by a loop antenna over a ground plane. The strength of the radiated field is directly proportional to the amplitude of the current, the effective loop area and the square of the frequency of the current [1].

As has already been discussed, the magnitude and high frequency content of the current is significantly lower in self-timed designs than in equivalent synchronous designs. Since communication in a self-timed design is local and there is no global clock signal, it is expected that the effective loop area of a self-timed design will be less than that of a synchronous design. This second point, however, requires further investigation and in this paper both loop areas are assumed to be equal to the core size.

Plots of the calculated field strengths for the self-timed and synchronous implementations at a measuring distance of 3m are shown in Figure 5. The dependence of EMI on the square of the frequency results in the synchronous design emitting radio-frequency energy an order of magnitude greater than the self-timed design at high frequencies. As expected, the EMI of the synchronous DSP shows peaks due to the clock harmonics while the EMI of the self-timed DSP is essentially white.

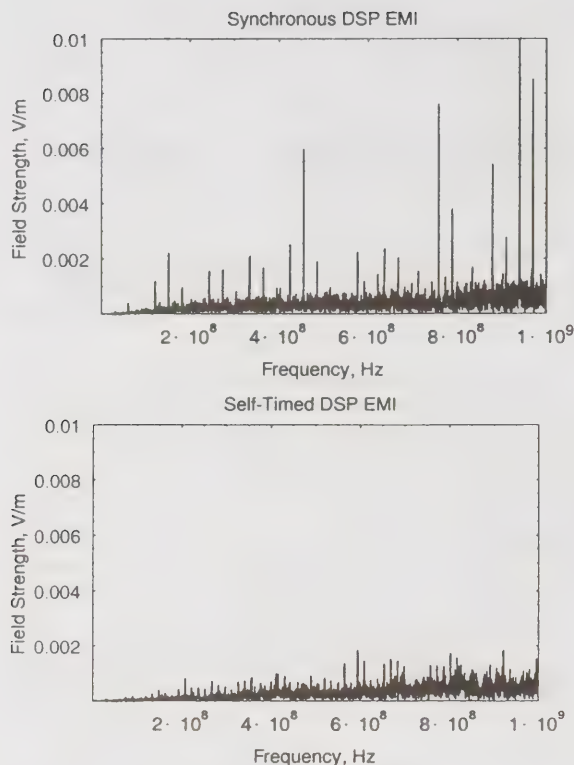


Figure 5. EMI Comparison

5. Conclusions

The paper has described several noise benefits of self-timed design. The results indicate that self-timed technology greatly reduces the high frequency components of the switched current. Since substrate coupling, packaging noise, and electromagnetic interference are all essentially high pass in nature, eliminating the high frequency components of the current greatly reduces the generated noise.

Substrate coupling in the synchronous design shows strong peaks at harmonics of the global clock. In the self-timed design the substrate coupling is broadband white noise of lower magnitude.

Packaging noise in the synchronous design causes large transients in the ground voltage near the edges of the global clock. No such transients are observed in the self-timed design.

EMI in the synchronous design shows strong peaks at harmonics of the global clock. In the self-timed design, the EMI is reduced by an order of magnitude and shows no peaks at high frequencies.

6. Acknowledgments

The development of the self-timed DSP core has been carried out under contract and with the co-operation of LG Semicon, Seoul, Korea. The authors are grateful for the assistance and support provided by LG Semicon.

7. References

- [1] Henry W. Ott, "Noise Reduction Techniques in Electronic Systems", p. 298, John Wiley & Sons, 1989.
- [2] Ranjit Gharpurey and Robert G. Meyer, "Modeling and analysis of substrate coupling in integrated circuits", *IEEE J. Solid State Circuits*, vol. 31, no. 3, pp. 344-353, Mar. 1996.
- [3] Nishath K. Verghese, David J. Allstot, Mark A. Wolfe, "Verification techniques for substrate coupling and their application to mixed-signal IC design", *IEEE J. Solid State Circuits*, vol. 31, no. 3, pp. 354-365, Mar. 1996.
- [4] B. R. Stanisic, N. K. Verghese, L. R. Carley, and D. J. Allstot, "Addressing substrate coupling in mixed-mode IC's: Simulation and power distribution synthesis," *IEEE J. Solid State Circuits*, vol. 29, no. 3, pp. 226-238, Mar. 1996.
- [5] P. Miliozzi, I. Vassiliou, E. Charbon, E. Malavasi, and A. Sangiovanni-Vincentelli, "Use of sensitivities and generalized substrate models in mixed-signal IC design," *Proc. IEEE/ACM Design Automation Conference*, Jun. 1996.
- [6] N.C. Paver, P. Day, C. Farnsworth, D.L. Jackson, W.A. Lien, J. Liu, "A Low-Power, Low Noise, Configurable Self-Timed DSP", *Proc. 4th Int. Symp. on Adv. Research in Asynchronous Circuits and Systems*, pp. 32-42, March 1998.
- [7] S.B. Furber, P. Day, J.D. Garside, J. Liu, N.C. Paver, S. Temple, "AMULET2e: An Asynchronous Embedded Controller," *Proc. of Advanced Research in Asynchronous Circuits and Systems*, Eindhoven, April 1997.
- [8] I.E. Sutherland, "Micropipelines," *Communications of the ACM*, vol. 32, no. 6, pp. 720-738, June 1989.
- [9] C.H. van Berkel, M. Rem, R.W.J.J. Saejis, "VLSI Programming," in *Proc. of ICCD'88*, pp. 152-156, 1988.
- [10] "GM5288 (V.34 Fax/Modem Data Pump)", Data Sheet, LG Semicon Co. Ltd., Seoul, Korea, 1996.
- [11] Ranjit Gharpurey, "Modeling and Analysis of Substrate Coupling in Integrated Circuits," Ph.D. thesis, University of California at Berkeley, 1995.

Wideband radio prototyping environment

Tommi Mäkeläinen, Timo Eriksson, Jarkko Posti, Risto Suoranta
Nokia Research Center

Tel +358 3 272 5749, fax +358 3 272 5935,
e-mail: tommi.makelainen@research.nokia.com

Abstract

Evaluation of the candidates for the next generation cellular wireless standards is currently going on widely. One critical aspect in the standardisation process is the definition of terrestrial physical layer interface. The physical layer principles along with resource allocation mechanism on Medium Access Control (MAC) together define how well the physical layer resources can be utilised. The prototyping environment presented in this paper has been built to analyse and validate effectiveness of the different design choices. As an example, the case considered is Wideband Time Division Multiple Access (W-TDMA) based air interface developed in the EU ACTS AC090 FRAMES project.

Key words: radio prototyping environment, flexible design flow, cellular link validation, validation of packet access services

Introduction

Validation of a candidate wireless communication system is time consuming task. Due to complexity of the task, it is important to design the system so that design parameters both on physical and MAC layers can be modified without complete re-design of the validator. In this paper, we describe a wideband radio prototyping environment where design parameters can be modified easily and the effects can be analysed and validated both off- and on-line with the real-time prototype system. As an example, the case considered is W-TDMA based air interface developed in the EU ACTS FRAMES. The architecture and main capabilities of the W-TDMA system are shown.

Keys for flexible prototyping

There are number of methods that can be used to ensure reasonable design iteration times. The ones used in our prototyping system are

1. simulator based design and implementation,
2. use of standard system components and interconnections,
3. configuration flexibility, centralised system control,
4. testing of analysis functions in simulator environment

Prototyping environment

Simulator based design and implementation in this case means the use of software simulators to first implement and functionally simulate major parts of the system before implementation on the final hardware. Figure 1 illustrates main components of the design environment.

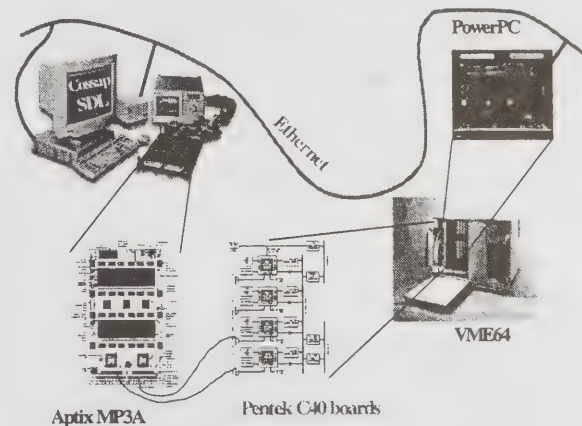


Figure 1: Design environment

Protocol and system control software has been implemented using Specification and Description Language (SDL) using Telelogic Inc. TAU-toolset, with some lower control functionality and protocol data unit processing implemented in C. Most of the baseband algorithms are written using generic C in Synopsys Cossap environment, while in addition

some custom coding with assembler/C, and VHDL has been used. Both the simulation environments support generation of C-language implementation down to target platform. The selected approach offers simulation and testing of the functionality in workstation environment, and a direct route down to actual hardware to iterate the design.

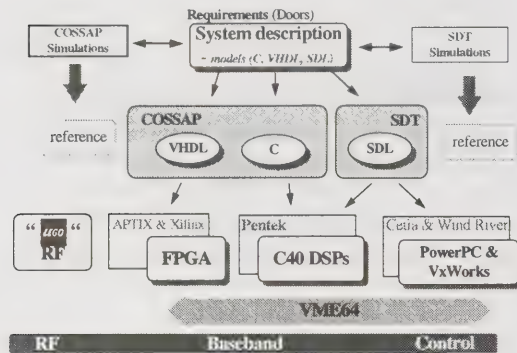


Figure 2: Overall design flow

To efficiently utilise flexibility provided by the simulator based design, platform must be based on standard system components and interconnections. Use of standard components facilitates use of code generation both for HW and SW. The standard interconnections allows easy modifiability of the target. Our system uses C40 communication links, and MIX and VME buses. User and control/signalling data flows are internally separated to ensure flexible modifiability and deterministic timing. This separation combined with the idea of centralized system control suits well for the selected design methodology, where user data processing and system control and signalling development is done on Cossap and SDT, respectively.

W-TDMA test scenario

The test case used is W-TDMA based air interface developed in the EU ACTS AC090 FRAMES. It used GSM frame period with 1/16 and 1/64 slots. Air interface supports both real-time (RT) and non real-time (NRT) connections. Detailed description of the W-TDMA parameters can be found in [1] and [2].

The platform consists of a mobile station (MS) and a base station (BS). Functionally both the MS and BS have been built to resemble each other. As the target of prototyping is to validate the concept, no product related design aspects, like power

consumption or cost, have been considered. A HP UNIX workstation, connected through Ethernet to both MS and BS, is used for centralized control and for off- and on-line analysis displays.

Internally a MS or BS is divided to three separate logical units, namely RF, baseband and service/control (CTRL).

RF Unit

The RF Unit is split into three sub-units, namely RF front-end, RF MIX board, and RF DSP board. RF front-end module upconverts the Tx intermediate frequency (IF) signal (generated by the Baseband Unit) into the final RF channel, amplifies it to the required level, and performs the transmission. RF front-end receives the RF signal located on a specific RF channel, amplifies and filters it, and downconverts to a fixed low intermediate frequency which can be processed by the BB Unit. RF MIX board sub-unit is responsible for controlling the RF front-end parameters like Tx/Rx frequencies, and Tx/Rx gains. In addition it generates reference clock for the RF front-end, CTRL unit, and sample + Tx/Rx gating clocks for the baseband unit. RF DSP board implements the interface between the VME, and MIX busses. It perform configuration information processing into a suitable format for the MIX board, and RF front-end.

RF unit is similar between the base, and mobile station applications. The only difference is on the MS side where the reference oscillator frequency can be adjusted by means of a voltage controlled oscillator.

The detail properties of RF unit are listed below:

Item	unit	value
Frequency ranges	MHz	1920-1980, 2110-2170
RX noise figure	dB	4
Channel spacing	MHz	1.6
RX AGC range	dB	80
Tx power control range	dB	50
Frequency switched speed (Tx and Rx)	us	1
Tx IF	MHz	14
Rx IF	MHz	30
Reference frequency	MHz	104

Baseband Unit

Baseband unit consists of several sub-modules, namely transmitter encoders, burst multiplexer, transmitter front-end, receiver front-end, burst demultiplexer, and receiver equalizers / decoders. All of these are physically implemented using multiple Texas Instrument's TMS320C40 DSP processors on Pentek's off-the-shelf boards, and still a few custom hardware boards. Figure 3 shows functional architecture of baseband unit.

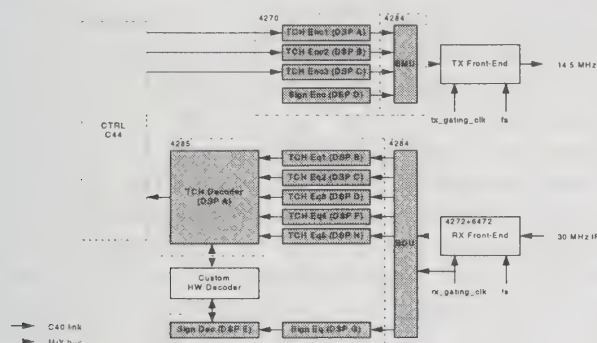


Figure 3: Baseband architecture

Most of the algorithms are written using generic C in Cossap environment, while in addition some custom coding with assembler/C, and VHDL must be performed. The baseband units both in BS and MS are similar, with the exception on MS side due to additional measurements for synchronising the radio link.

On the transmitter side there are three traffic, and one signalling channel encoder responsible for error detection / protection coding, interleaving, and burst building. On the receiver side the IF signal is first sub-sampled with A/D converter, after which a receiver board is utilized to perform down-conversion (to baseband), and reception filtering. Received bursts are routed to five traffic, or one signalling equalizer via burst demultiplexer unit. Equalizers perform multipath effect cancellation by means of a decision feedback equalizer. Following this the soft bits are deinterleaved and weighted according to burst SIR, and in case of ready packet convolutionally decoded. CRC is finally performed for the packet.

Control Unit

Control Unit implements centralised control of physical layer and MAC functionality to allocate resources for circuit switched and packet access connections, simple link access control, and simple

network layer functionality to manage user data connections. The control of Baseband and RF is done on a TDMA-frame basis.

A circuit switched connection is provided as a comparison system to evaluate performance of the packet switched connection. Packet switched service is pure Non real-time service, so no maximum delay figures are offered. Selective receiver controlled hybrid-II ARQ is used as error protection scheme for packet access connections. Most of the functionality have been implemented using SDL-language to specify the functionality. Figure 5 illustrates protocol entities implemented. Acronyms used in Figure are Radio Link Control (RLC), Logical Link Control (LLC), Radio Bearer Control (RBC) and Internet Protocol (IP).

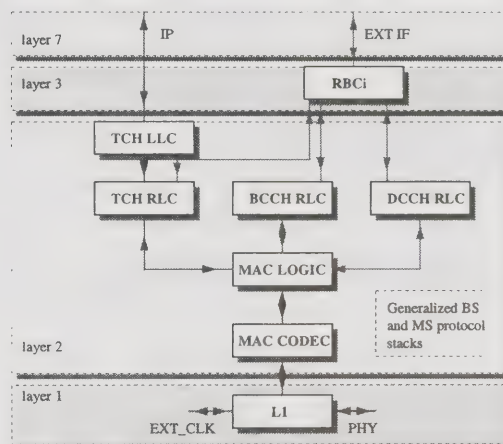


Figure 4: SDL design flow and protocol stack implemented in SDL

Figure 4 presents the design flow for the modules implemented in SDL. As a starting point, Message Sequence Charts (MSCs) are used to identify flow of events for all the system behavior. After completing MSCs, SDL is used to design the protocol using parallel extended state machines. After workstation based simulation, validated against initial MSCs, is satisfactory, target processor C-language code is generated from the SDL specification.

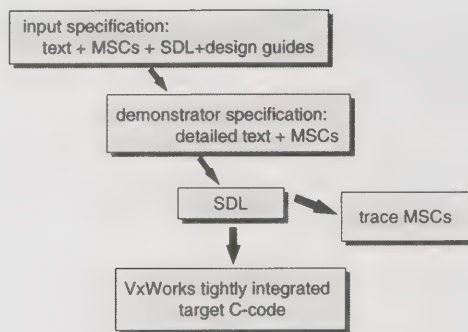


Figure 5: Protocol design flow

The actual PDU processing and Physical layer functionality have been implemented with hand-coded C-language routines. User data connection provided is TCP/IP over Ethernet. All common IP services can be used, e.g. web-browsing.

Analysis Unit

Support for system analysis is provided by implementing measurement functions directly into the data path and control flow processes. Analysis data is collected from all the protocol layers, from the layer 1 to layer 7. All the measurements are first tested in simulation environments, and then validated on the real platform. The selected method ensures that adding analysis procedures do not change the behaviour of the system. System provides support for storing the main system parameter on a physical layer frame basis to a hard disc for post processing and analysis, and for a set of parameters real-time display is available. A UNIX workstation, connected through Ethernet to both MS and BS, is used for centralized control and for off- and on-line analysis displays. On-line analysis values can be traced either as discrete values or as averaged values over time.

Conclusions

The presented prototyping environment is a powerful environment to analyse the performance of W-TDMA based radio links. It facilitates easy

modifiability and seamless design flow for functional prototyping purposes. It is relatively easy to modify for other TDMA based wireless systems.

Acknowledgements

This work has been partially performed in the framework of the project ACTS AC090 FRAMES, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues from Siemens AG, Roke Manor Research Limited, Ericsson Radio Systems AB, Technical University of Delft, University of Oulu, France Telecom CNET, CSEM-Centre Suisse d'Électronique et de Microtechnique SA, Swiss Federal Institute of Technology Zurich, University of Kaiserslautern, Chalmers University of Technology, The Royal Institute of Technology, Instituto Superior Técnico and Integracion y Sistemas de Medida S.A.

The authors would like to acknowledge the following persons that have participated in the work described in this paper: Jouni Siirtola, Jukka Tapaninen, Hans-Otto Scheck, Sari Suomi, Olli Koskinen, Martti Virtanen, Cui Wei, Tiina Toivanen and Juri Sipilä.

References

- [1] A. Klein, R. Pirhonen, J. Sköld, and R. Suoranta, "FRAMES multiple access Mode 1 - wideband TDMA with and without spreading", in Proc. Int. Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, Sep. 1997, pp. 37 - 41.
- [2] E. Nikula, A. Lappeteläinen, J. P. Castro, "Performance of FRAMES multiple access mode 1 without spreading - WB-TDMA", in Proc. 3rd ACTS Mobile Communication Summit, Rhodes, Greece, June 1998, pp. 327 - 332.

Digital Single Side Band with Pilot Symbol Assisted Modulation in Mobile Radio Channels

Seungwon Kim,
Leon W. Couch II

Dept. of Electrical and Computer Engineering
University of Florida
P.O. Box 116200

227 Larsen Hall, Gainesville, FL 32611-6200
352-392-8590, 352-392-0044(fax), swkim@grove.ufl.edu

Abstract

This paper presents the analysis of digital single side band(DSSB) modulation technique with pilot symbol assisted modulation(PSAM). As a measure of difficulty of linearization, the dynamic range, defined as the ratio of maximum to minimum instantaneous powers is commonly used. DSSB modulation is shown to provide the much reduced dynamic range from 5.5dB to 8.9dB whereas the dynamic range of QPSK is infinite and the dynamic range of OQPSK is around 11dB. With DSSB, therefore, we can adopt more efficient amplifier with much simplified linearizer design. Besides, the proposed DSSB system does not need very sharp cutoff filtering and only sloped filtering, such as a raised cosine filtering, can be used.

1. Introduction

The efficiency of channel usage could be improved by the use of schemes such as dynamic channel allocation, cellular radio and some forms of modulation such as narrow-band single side band(SSB). SSB modulation provides effective utilization of the available frequency spectrum with relatively low power. Although SSB has bandwidth efficient advantage, SSB system performance in a fading channel is poor. The use of a pilot tone for synchronization of the SSB receiver and to mitigate the effects of fading has been explored by several authors[1] -[2]. Among these pilot tone assisted modulation (PTAM) methods, transparent tone in band (TTIB)[3] is the most useful one. However, this method requires complicated signal processing such as frequency shift and filtering. TTIB results in an increased peak-to-average power ratio compared to that of pilot symbol assisted modulation (PSAM)[4]. With PSAM, A known pilot symbol into the data is periodically transmitted to analyze fading distortion. Both PSAM and PTAM mitigate the effects of multipath fading, but PSAM has more advantages than PTAM[4].

In this paper, we define DSSB as the digital SSB modulation technique using a pulse shaping filter at the transmitter and a low pass matched filter at the receiver. DSSB with PSAM will be

investigated under Rayleigh fading channel conditions.

2. DSSB PSAM System

2.1 System block diagram

The block diagram of the DSSB PSAM system is shown in Fig. 1. Known pilot symbols are inserted periodically into the every frame of length N symbols as shown in Fig. 2. The transmit low-pass filter is a square root raised cosine filter as shown later. The modulated DSSB signal is transmitted in the usual way over a channel characterized by flat fading and AWGN. The demodulated signals are sampled at the symbol rate $1/T_s$ and the frame rate $1/T_F = 1/NT_s$. It is assumed that this timing is regenerated perfectly. The samples at kT_F correspond to the pilot values out of the receiver low pass filter. These pilot values are used to estimate the channel state.

2.2 DSSB Modulation

We assume that here the phase shift method is used for the proposed DSSB system as shown in Fig. 1. In this case, we require only the square root raised cosine filter and the Hilbert transform of the square root raised cosine filter.

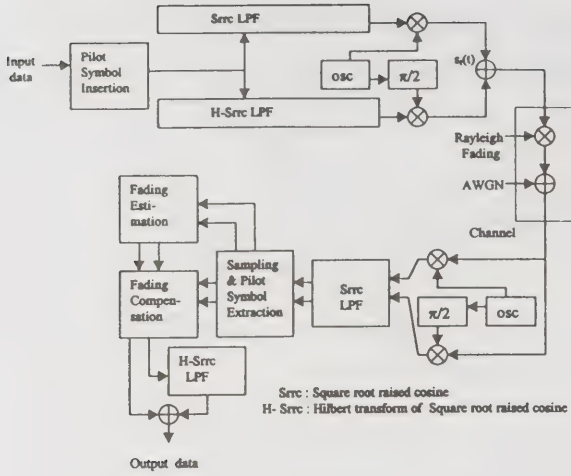


Fig. 1. DSSB PSAM system

The output of the Hilbert transform of the square root raised cosine filter will be truncated by using Gaussian window to decrease both ISI of adjacent symbol and peak power. The modulated signal $s_t(t)$ at the transmitter output is given by

$$s_t(t) = A_m(t)\cos(2\pi f_c t) \pm A_{m_h}(t)\sin(2\pi f_c t) \quad (1)$$

where A is the amplitude of input data $d(t)$ that corresponds to binary digit 1 (or $-A$ for 0), the (+) sign used for LSSB, (-) sign is used for USSB, $m(t)$ is the normalized square root raised cosine pulses and $m_h(t)$ denotes the Hilbert transform of $m(t)$ as shown in Fig. 3.

$$m(t) = [\sin((1-r)\pi t/T_s) + (4rt/T_s)(\cos((1+r)\pi t/T_s))] / (\pi t/T_s)(1 - (4rt/T_s)^2) \quad (2)$$

$$m_h(t) = m(t) * h(t) = 1/(\pi t/T_s) + (4r/\pi)[\text{sinc}((1+r)t/T_s)(1+r)(\pi t/T_s) - \cos((1-r)\pi t/T_s)/(4rt/T_s)] / (1 - (4rt/T_s)^2) \quad (3)$$

$$\text{where } h(t) = 1/\pi t \quad (4)$$

and r is the roll off factor ranges between 0 and 1. If we assume USSB is used to transmit the modulated signal $s_t(t)$, this can also be represented as

$$s_t(t) = A\text{Re}[z_T(t)\exp(j2\pi f_c t)] \quad (5)$$

$$\text{where } z_T(t) = m(t) + jm_h(t) \quad (6)$$

is the complex baseband signal of the transmitted signal which is shown in Fig. 3.

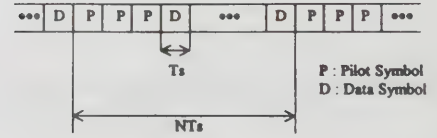


Fig. 2. Transmitted frame structure

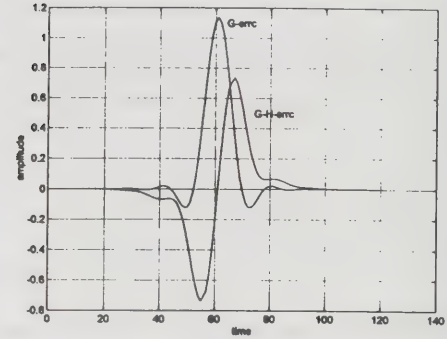


Fig. 3. The square root raised cosine pulse with Gaussian window('G-src') and the Hilbert transform of the square root raised cosine pulse with Gaussian window('G-H-src') with roll off factor $r = 0.35$

2.3 The problem using only one symbol as a pilot symbol

Pilot symbol must provides the receiver with an explicit amplitude and phase reference for detection. However, because of the inter symbol interference(ISI) of the Hilbert transform of the square root raised cosine pulse, there is an amplitude and phase ambiguity in case of using only one symbol as a pilot symbol as shown in Fig. 4. To greatly reduce this ISI, we add more adjacent symbols as pilot symbols. Hence, we use three adjacent pilot symbols in one frame and amplitude and phase estimation of the fading will be made at the middle point in the middle pilot symbol. This is shown in Fig. 5.

2.4 Demodulation

The received signal is expressed by

$$s_r(t) = \text{Re}[Ac(t)z_R(t) + n_c(t)] \quad (7)$$

where $c(t)$ represents the complex Gaussian fading which has a complex zero mean.

$$c(t) = \alpha(t)e^{j\theta(t)} \quad (8)$$

The power spectrum of $c(t)$ is given by[5]

$$C(f) = \sigma_c^2 / 2\pi\sqrt{(f_D^2 - f^2)} \quad (9)$$

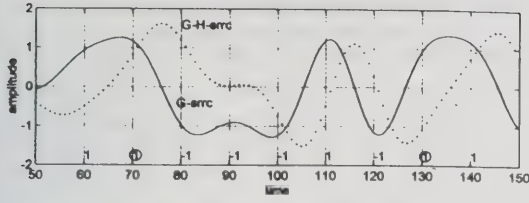


Fig. 4. Amplitude and phase ambiguity in case of one bit pilot symbol

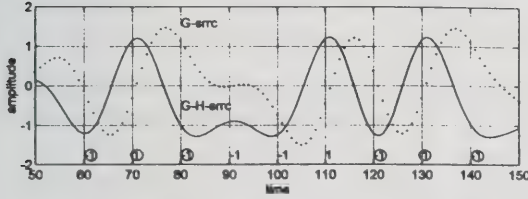


Fig. 5. Explicit amplitude and phase reference in case of three pilot symbols

where f_D indicate the maximum Doppler frequency. The demodulated and low pass filtered complex signal can be obtained by,

$$x(t) = [Ac(t)z_R(t) + n(t)]\exp(-j2\pi f_{off}t) \quad (10)$$

Here, $z_R(t)$ is the signal component of the received complex baseband signal and $n(t)$ is AWGN with power spectral density $N_0/2$ in both real and imaginary components. The receiver only works if $2(f_D + f_{off}) \ll D/N$.

2.5 Sampling and Pilot symbol extraction

The sampling time of demodulated complex baseband signals $x(t)$ is $1/T_s$ and the frame rate is given by $1/T_F = 1/NT_s$. The pilot symbols are sampled at kT_F and uncompensated received signals are sampled at $(k + m/N)T_F$, $m = 0, 1, \dots, (N-1)$.

2.6 Fading estimation and compensation

Normally, the pilot symbol would be randomized to avoid transmission of a tone, and the receiver would make appropriate corrections based on its knowledge of the transmitted pilot values. However, in the following analysis of fade compensation, it is assumed for simplicity that a constant pilot sequence have value $d' = A$ and data symbols are $d(k) = \pm A$. The k th pilot symbol at $t = kT_F$ is given

$$x(k) = [Ac(k)z_R(k) + n(k)]\exp(-j2\pi f_{off}kT_F) \quad (11)$$

The fading channel estimation $\hat{c}(k)$ of $c(t)$ at $t = kT_F$ is obtained by dividing $x(k)$ by the corresponding transmitted pilot symbol. Hence,

$$\hat{c}(k) = x(k)/d' \quad (12)$$

The other points fading distortion at $t = (k + m/N)T_F$ is given by Gaussian interpolation of the fading channel estimates $\hat{c}(k)$.

$$\hat{c}(k + m/N) = Q_{-1}(m/N)\hat{c}(k-1) + Q_0(m/N)\hat{c}(k) + Q_1(m/N)\hat{c}(k+1) \quad m = 0, 1, 2, \dots, (N-1) \quad (13)$$

where coefficients Q_i is given in [6]. In case of first order and zeroth order interpolation, the coefficients Q_i is obtained as, $Q_{-1} = 0$, $Q_0 = 1 - m/N$, $Q_1 = m/N$. The complex baseband samples are compensated by calculating (14).

$$\hat{z}(k + m/N) = x(k + m/N) / \hat{c}(k + m/N) \quad m = 0, 1, 2, \dots, (N-1) \quad (14)$$

Hence, the decision input is given by adding both the real components of $\hat{z}(k + m/N)$.

3. Computer simulation performance

3.1 Dynamic range

Nonlinear amplification of the zero-crossings can bring back the filtered sidelobes. Hence, linear amplifiers which are less efficient should be used to prevent the spectral widening. If the signal is limited to an annular region over which the amplifier non-linearity is moderate, it is easier to linearize these amplifiers. As a measure of difficulty of linearization, the dynamic range, defined as the ratio of maximum to minimum instantaneous powers is commonly used. Table 1 shows the dynamic range for DSSB, QPSK and OQPSK. QPSK signal envelope pass through zero due to phase shift of π radians. The dynamic range is infinite and require more effort in the design of a linearizer, or more expensive and less efficient amplifiers. OQPSK signal envelope does not go to zero, since π phase transitions have been removed. The envelope variations are less than that of QPSK. The dynamic range of OQPSK is around 11dB whereas the dynamic range of DSSB is from 5.5 to 8.9dB. With DSSB, therefore, we can take advantage of the much reduced dynamic range and adopt Class C or Class AB amplifiers with much simplified linearizer design.

3.2 Peak to average power ratio

Due to Hilbert transform of the square root raised cosine pulse, the pulse shaped DSSB has less envelope fluctuation than that of QPSK. However, the worst case peak power of the DSSB occurs for the case of a long stream of ones and zeros. This worst case peak power can be eliminated by truncating and windowing the pulse shape filter. Table 1 also shows peak to average power ratios for DSSB, QPSK and OQPSK.

3.3 Spectral occupancy

The proposed DSSB requires windowing to reduce ISI and peak power. This causes the increase of spectral occupancy of DSSB. The PSD of DSSB with Gaussian windowing is shown in Fig. 6. As shown in Fig. 6, the PSD of the proposed DSSB looks like that of vestigial side band (VSB). When the baseband signal contains significant components at extremely low frequencies, the use of SSB modulation is inappropriate for the transmission of such baseband signals due to the difficulty of obtaining the very sharp cutoff filtering needed. That is the reason why SSB can not be used with analog TV. But the proposed DSSB does not need the very sharp cutoff filtering. The gradual sloped filtering, such as raised cosine filtering, can be used for the proposed DSSB.

3.4 BER performance under Rayleigh fading

Fig. 7 shows the BER performance for several values of normalized Doppler frequency. The transmission rate is 40kbaud and this Doppler frequency ranges from 40Hz to 400Hz. It is shown that the performance degradation is within 2dB of theoretical BER value under Rayleigh fading.

References

- [1] R. Wells, "SSB for VHF mobile radio at 5Khz channel spacing", IERE conf., Proc. Radio receivers and associated syst., South Hampton England, July 1978, pp. 29-36
- [2] B.B. Lusignan, "single-sideband transmissions for land mobile radio", IEEE spectrum, pp 33-37, July 1978
- [3] J.P. McGeehan and A.J. Bateman, "Phase locked transparent tone-in-band(TTIB) : A new spectrum configuration particularly suited to the transmission of data over SSB mobile radio networks", IEEE Trans. Commun., vol. com-32 pp 81-87, Jan. 1984
- [4] Jakes K cavers and Maria Liao, "A comparison of pilot tone and pilot symbol techniques for Digital Mobile communication", IEEE Global Telecommun., conf., pp 915-921, Dec., 1992.

[5] W.C. Jakes, Jr., Microwave Mobile Communications, New York Wiley, 1974.

[6] S.Sampe and T. Sunaga, "Rayleigh fading compensation method for 16 QAM in digital land mobile radio channels", in Proc., IEEE Veh., Technol., conf., San Francisco, CA, pp. 640-646. May 1989.

	DSSB			QPSK	OQPSK
Time span	$\pm 6T_s$	$\pm 10T_s$	$\pm 12T_s$	$\pm 6T_s$	$\pm 6T_s$
Peak Power	2.8301	4.0244	4.4994	5.00	4.60
Minimum Power	0.8028	0.6729	0.5790	0	0.38
Average Power	1.6689	1.8202	1.8560	2.00	2.00
Dynamic Range(dB)	5.47	7.76	8.90	∞	10.83
Peak to Average Power Ratio(dB)	2.29	3.44	3.84	3.97	3.62

Table 1. Dynamic range and Peak to Average Power Ratio for DSSB, QPSK and OQPSK with the parameter of Time span and roll off factor $r = 0.35$

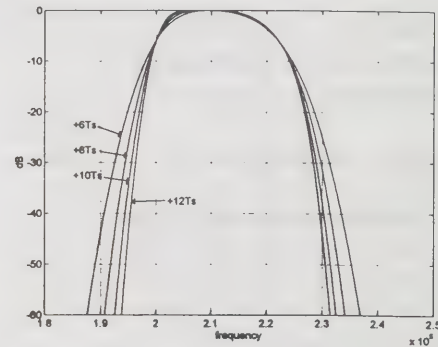


Fig. 6. PSD of DSSB with Gaussian window with symbol rate $D = 40000$, roll off factor $r = 0.35$ and carrier frequency $f_c = 2 \times 10^5$

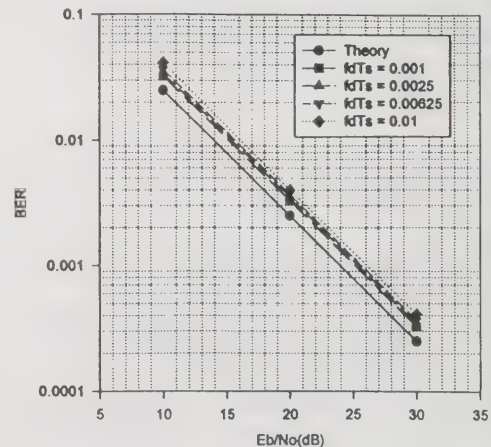


Fig. 7. BER of DSSB PSAM under Rayleigh fading with $r = 0.35$, Time span $= \pm 12T_s$, 1st Gaussian interpolation and Frame length $N = 20$

Characterization of Bunny-Ear Antennas for Wireless Basestation Applications

Marc C. Greenberg and Kathleen L. Virga
University of Arizona
Tucson, AZ 85721

520-626-2538, 520-621-2999 (fax), marcg@ece.arizona.edu

Abstract

The bunny-ear antenna (BEA) which has use for wireless basestation applications, mobile satellite communications, and radar imaging is characterized in detail. Whether the BEA is used as an isolated antenna or in an array, it needs to be well understood. A design methodology for the BEA and examples are presented. To discuss the effects from removing metalization near the feed, antenna surface currents are examined. The antenna's input impedance is examined and compared to the well-known Vivaldi antenna. The performance parameters return loss, frequency dependence of beamwidth, and far-field patterns are also investigated and compared with those of the Vivaldi.

Key words: Tapered slot antenna, Vivaldi aerial, Bunny-Ear antenna

Introduction

There is wide interest for antennas that are simple low-profile structures with broadband characteristics for wireless basestation applications, mobile satellite communications, and radar imaging. For basestations, omnidirectional antennas are usually the preferred choice, but the gain that would be the same in all directions in the azimuthal plane over a flat terrain is no longer omnidirectional in the case of rough terrain. The rough terrain can reshape the far-field pattern causing unwanted interference in certain directions. For locations with such terrain, a directional antenna is sometimes the preferred choice [1]. Desired features for such antennas include simple and low-cost fabrication. A modified, tapered slotline antenna introduced in [2] is a good candidate for meeting these goals. Such an antenna can be used as the base element in an array [2] or as a stand-alone antenna. In this paper, the performance and design characteristics of the isolated bunny-ear antenna for wireless basestation applications are investigated and compared to those of the Vivaldi antenna.

Antenna Geometry and Simulation Methodology

The BEA is a modified Vivaldi radiator, a type of tapered slot antenna. Like a Vivaldi radiator, the BEA is an endfire antenna, where the amount of flare of the inner edge of each conducting arm is described by an exponential. The major difference between the bunny-ear and the Vivaldi is that the conductors are narrower

at the feed than at the radiating end of the BEA. The outer edge of each conducting arm on the BEA loosely follows the slot taper and can also be described by an exponential. Fig. 1 shows a basic geometry for the BEA. As is with the Vivaldi, the mechanism behind the antenna's operation is its broadband matching transition from the characteristic impedance at the feed to the characteristic impedance of free space at the slotline's far end. The outer edge of the conductor is tapered to reduce the shunt capacitance between one arm of the slotline and a coaxial mounting block used for feeding the antenna with a coaxial line. The shunt capacitance created shorts out the lower frequencies of the antenna at the feed. For an array, the taper of the outer conductor of the BEA keeps the feeds of adjacent element isolated.

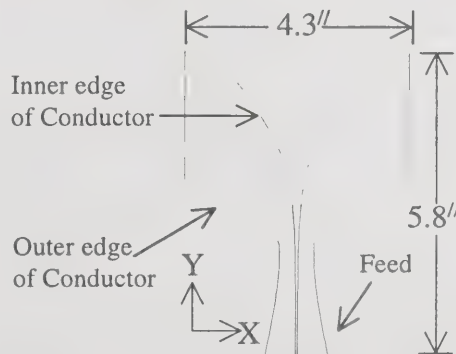


Fig 1 An example of a bunny-ear antenna.

For an air-backed Vivaldi, only the slot width determines the characteristic impedance of the

transmission line, where for a BEA the finite width of the conductors is also a contributing factor to the characteristic impedance of the line. A program based on the theory developed in [3] was written for calculating the characteristic impedance of a slotline with finite width conductors. The program can be used to generate impedance profiles instrumental in the design of tapered slot antennas (TSAs).

To evaluate the attributes and performance of the BEA the method of moments (MM) was employed. For the surface patch MM, the antenna is divided into a grid of triangular patches [4]. Applying the electric field integral equation (EFIE) to the infinitely thin, perfectly conducting surface of an antenna generates a matrix equation. The RWG basis functions are the basis functions used in the EFIE. The solution to the matrix equation is the surface currents on the antenna. From the surface currents, the input impedance and radiation patterns can be calculated. Initial BEA simulations were verified by measurement, and were in good agreement with one another. From convergence tests, it has been found that in order to achieve accurate predictions in input impedance the mesh should be designed fine enough to make the largest edge length less than $\lambda/8$ at the highest frequency to be simulated.

Design Considerations

For the air-backed Vivaldi, the overall length and width of the antenna determine the taper. A BEA with the same overall dimensions as a Vivaldi will have the same inner slot taper as the Vivaldi. Since the slot taper is determined from the width and length of the antenna, it is necessary to set some criteria for those dimensions. For a TSA to behave as a travelling wave antenna and exhibit nearly constant input impedance, the antenna should be more than three wavelengths long [5]. The width of the slot determines the lowest frequencies that radiate. Specifically the lowest frequencies that will radiate have wavelengths around twice the widest portion of the slot. The BEA has the extra design parameters of how wide the conductor is at the feed and how quickly the outer conductor edge tapers.

Fig 2 shows four antenna configurations explored in this work. Each antenna is assumed to be mounted on a substrate with $\epsilon_r=1.0$. For each of the four antennas, the inner flare width is described by $s = .03 \cdot \exp(0.615 \cdot z)$ where the units are inches and the flare width at the feed is 30 mils. Case (a) is a Vivaldi antenna from which the other three cases are derived. Cases (b), (c), and (d) have the outer conductor edge tapered by the function $w = 1 \cdot \exp(kw \cdot z^2)$ and a width of one inch at the feed. Case (b) has the edge taper constant, kw , set to

$0.3/\text{in}^2$. Cases (c) and (d) have edge taper constants of $0.065/\text{in}^2$ and $0.025/\text{in}^2$, respectively. In case (a), $kw=0$.

The distance has been squared in the conductor taper function to exaggerate the difference between case (a) and the other three cases. If the distance is not squared then the conductor tapers quickly enough to the side as to make the antenna look very similar to a Vivaldi. This is true since the feed of the BEAs presented here is wide with respect to the overall width, and therefore tapers rapidly into a Vivaldi.

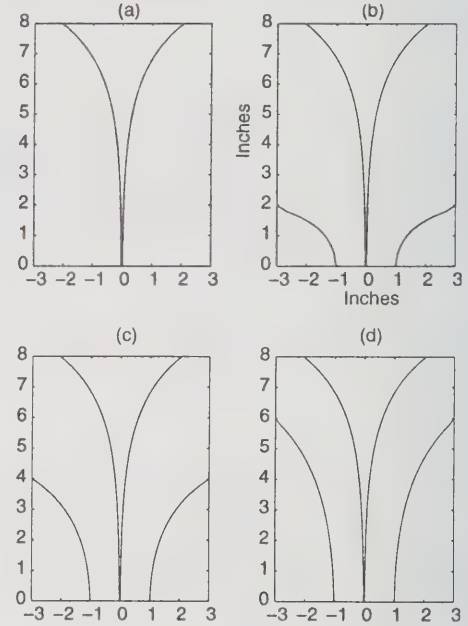


Fig 2 The four geometries explored. Case (a) is a Vivaldi so $kw=0/\text{in}^2$; case (b) has a conductor taper of $kw=0.3/\text{in}^2$. Case (c) has $kw=0.065/\text{in}^2$ and Case (d) has $kw=0.025/\text{in}^2$.

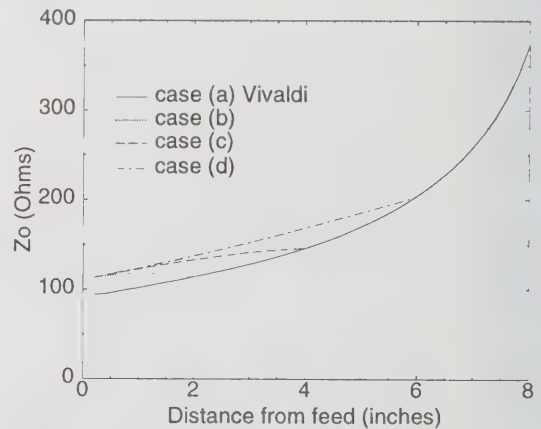


Fig 3 The impedance profiles for cases (a)–(d).

Fig. 3 shows an overlay of the slotline impedance profiles for the four cases. Note that the antennas all have very similar profiles. This is consistent since all

four designs have the same slot taper and terminate with the same dimensions. The difference is how far from the feed along the slotline before the slotline impedance is identical to the Vivaldi's.

Antenna Performance

Figs 4 and 5 show the antenna currents at 6 GHz for case (a) and case (d), respectively. The way the currents set up on the two antennas gives some insight into how their performance may differ. Fig 4 shows the Vivaldi antenna. It is clear from the figure that there is a standing wave pattern evident from the set of dark and light regions alternating in a line. As expected, the currents are most intense along the slot edges where the electric field is greatest. Fig 5 shows the BEA with the largest removal of metalization. Again the currents are most intense on the slot edges, but the standing wave pattern does not seem as well defined. This may lead to the hypothesis that removing metalization significantly degrades the antenna performance. This turns out not to be the case.

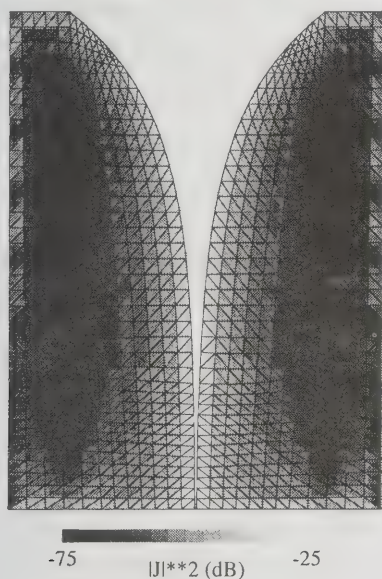


Fig 4 The current distribution for case (a) at 6 GHz.

Fig 6 shows the input resistance for the configurations presented. The input resistance for the BEAs stabilizes around 200 ohms as frequency is increased. It is unclear whether the Vivaldi's resistance stabilizes at 225 ohms or whether it would continue to climb as frequency is increased. Due to computer ram limitations in the simulation program, higher frequencies were not explored. One method to reduce the antenna's input resistance without increasing the length would be to mount the antenna on a dielectric substrate.

Fig 7 shows the input reactance for the different geometries. The reactance is a little high with a strong inductive component. For all cases the reactance tends toward values between 75 and 100 ohms.

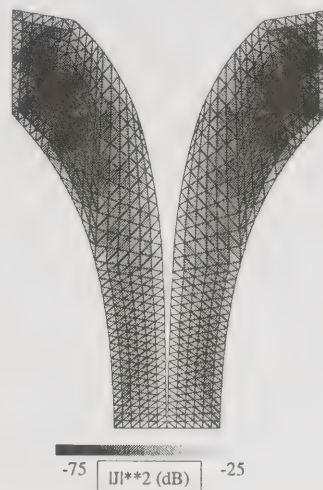


Fig 5 The current distribution for case (d) at 6 GHz.

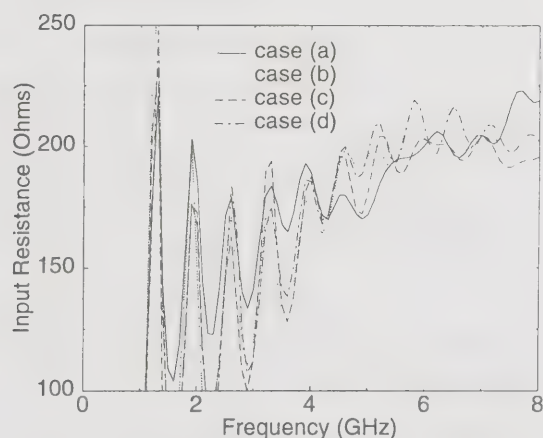


Fig 6 Input resistance vs. frequency for cases (a)—(d).

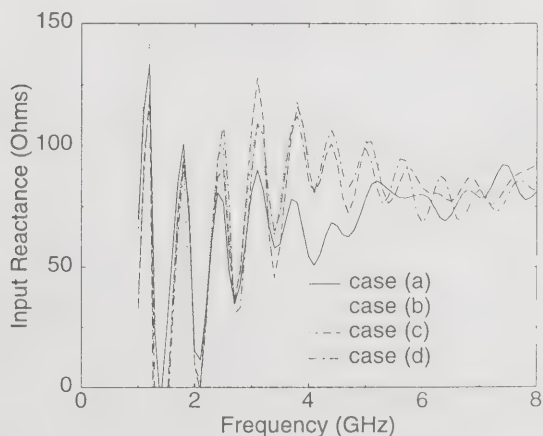


Fig 7 Input reactance vs. frequency for cases (a)—(d).

Fig 8 shows the return loss with respect to 200 ohms for the four cases. The antennas all have a first resonance near 1.4 GHz, which is nearly the frequency that corresponds to twice the maximum slot width—the condition stated in the Design Considerations section. It can be seen that the BEAs have a higher return loss than the Vivaldi does. This is due to higher reactance that is caused by the outer conductor taper.

The beamwidth of a Vivaldi antenna is theoretically independent of frequency [6]. Fig 9 shows the HPBW with respect to frequency from 1 to 10 GHz. For frequencies above 2.5 GHz the beamwidth does become relatively constant for all four cases. As more metalization is removed; that is, the conductor edge taper becomes a more dominant part of the antenna; the beamwidth varies less with frequency. Fig 10 shows a graph of relative power for each antenna at 6 GHz. The coordinate system is that shown in Fig 1. The relative power is only dependent upon the E_ϕ component. The Vivaldi has the third widest beamwidth, but tapers to -10 dB the fastest.

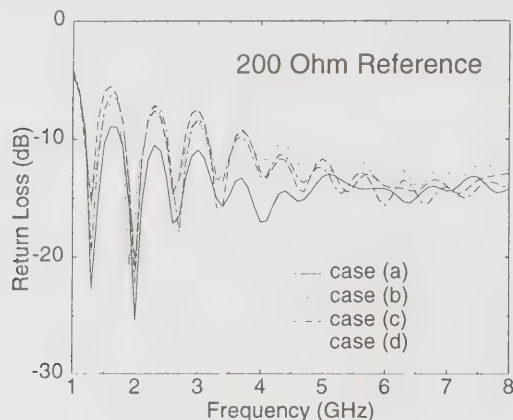


Fig 8 Return loss vs. frequency for cases (a)—(d).

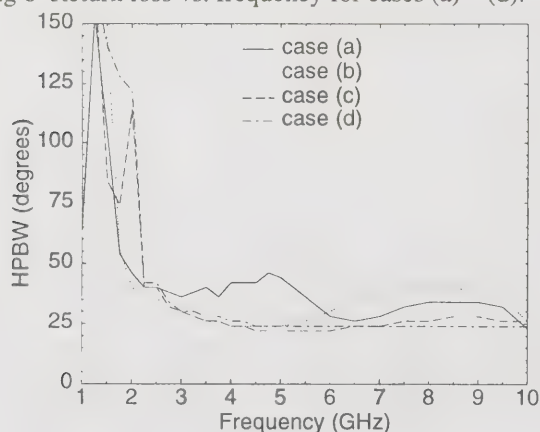


Fig 9 HPBW vs. frequency for cases (a)—(d).

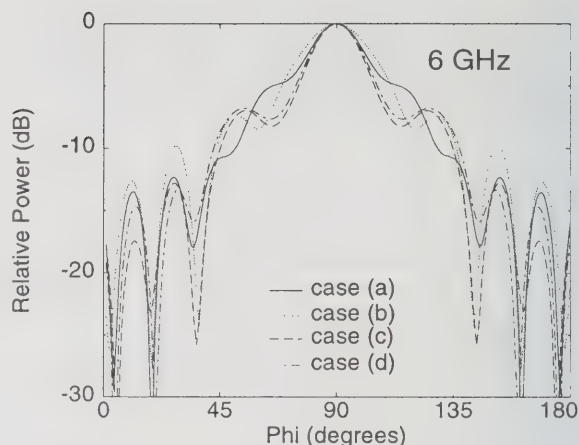


Fig 10 Far-field patterns at 6 GHz for cases (a)—(d).

Conclusion

A bunny-ear antenna derived from the Vivaldi antenna has been characterized. By tapering the conductor the designer has more control over the performance of the antenna. The BEA may be able to give superior frequency independence of beamwidth over the Vivaldi without degradation of the antenna pattern. A method to reduce the inductive component of the input impedance still needs to be explored.

Acknowledgement

The authors want to acknowledge the assistance of J. J. Lee and Stan Livingston for providing the slotline impedance program.

References

- [1] W. C. Y. Lee, *Mobile Communications Design Fundamentals*, John Wiley & Sons, New York, 1993.
- [2] J.J. Lee, S. Livingston, "Wide Band Bunny-Ear Radiating Element," *1993 IEEE Ant. and Prop. Society International Symposium*, pp. 1604-1607.
- [3] J.J. Lee, "Slotline Impedance," *IEEE Trans. Microwave Theory and Tech.*, Vol.39, No.4, Apr. 1991, pp. 666-672.
- [4] K. L. Virga and Y. Rahmat-Samii, "Low-Profile Enhanced-Bandwidth PIFA Antennas for Wireless Communications Packaging," *IEEE Trans. Microwave Theory and Tech.*, Vol.45, No.10, Oct. 1997, pp. 1879-1888.
- [5] D. H. Schaubert, "Endfire tapered slot antenna characteristics," *IEE Conf. Publ. Sixth International Conf. On Ant. and Prop.*, Coventry, England, Apr. 1989.
- [6] E. Gazit, "Improved design of the Vivaldi antenna," *IEE Proc.*, Vol.135, Pt.H, No.2, Apr. 1988, pp. 89-92.

EFFECT OF TILTED BASE-STATION ANTENNAS ON COCHANNEL INTERFERENCE REDUCTION IN CELLULAR COMMUNICATIONS

TING-KAI QIN XIN ZHANG & YONG CHENG
NANJING UNIVERSITY OF POSTS & TELECOMMUNICATIONS
NANJING 210003 CHINA
TEL +86-25-3492406 FAX +86-25-3492349
Email qintk@em.njupt.edu.cn xinzx@yahoo.com

Abstract

The effect of tilted base-station antennas on cochannel interference reduction is studied analytically in this paper, a work-site testing approach of antenna installation is proposed and better cochannel interference reduction can be expected by applying the new approach.

Key words: cellular communication, antenna, cochannel interference reduction

Introduction

Since the rapid development of mobile communications, cells in a cellular system tend to be much denser, and the frequency reuse ratio grows up tremendously. Therefore the reduction of cochannel interference between cellular cochannel sites becomes a serious problem which has to be solved. In this paper, we analyze the cause of tilted base-station antennas being of benefit to cochannel interference reduction, and conclude that if more attention is paid to the installation of antennas, more reduction of cochannel interference between the cells can be obtained besides designing a cellular system carefully.

Analysis

For a clear illustration, firstly we imagine a simplified case that Antennas A' and B' of two cochannel stations A and B respectively are faced to each other, and the antenna have

equal height h_t , and they emit the same power. A flat terrain between Stations A and B is also assumed. A mobile unit C communicates with Station A , and is interfered by the cochannel station B . The carrier to interference ratio of the mobile unit C can be written as[1]

$$\frac{C}{I} = \left(\frac{B'C}{A'C} \right)^4 \left[\frac{F(\theta_A)}{F(\theta_B)} \right]^2 \quad (1)$$

where

$$F(\theta) = \frac{\cos(2\pi \sin \theta)}{\cos \theta} \cdot \frac{\sin(2\pi \sin \theta)}{\sin\left(\frac{\pi}{2} \sin \theta\right)} \cdot \sin\left(\frac{\pi}{2} \cos \theta\right) \quad (2)$$

which is the pattern factor of antenna. θ_A and θ_B in (1) are the angles from the axes of the antenna main lobes to the lines $A'C$ and $B'C$ respectively, as shown in Fig.1.

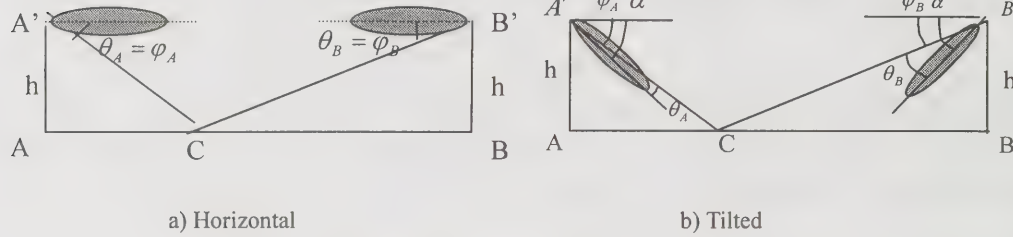


Fig. 1. The direction of antenna main lobes.

Assuming φ_A and φ_B are the angles of horizon to the lines $A'C$ and $B'C$ respectively, we have

$$\begin{aligned}\varphi_A &= \tan^{-1}\left(\frac{h}{AC}\right) \\ \varphi_B &= \tan^{-1}\left(\frac{h}{BC}\right)\end{aligned}\quad (3)$$

and

$$\begin{aligned}A'C &= \frac{AC}{\cos \varphi_A} \\ B'C &= \frac{BC}{\cos \varphi_B}\end{aligned}\quad (4)$$

When the axes of the main lobes are parallel to horizon, i.e., $\theta_A = \varphi_A$ and $\theta_B = \varphi_B$, it is found in (2) that $\theta_B < \theta_A$ produces $F(\theta_B) > F(\theta_A)$, which means the pattern factor of Antenna B' to the unit C greater than that of Antenna A' . In (1), we find that $(B'C/A'C)^4 \gg 1$ and

$[F(\theta_A)/F(\theta_B)]^2 < 1$, which leads to $\frac{C}{I}$ less than

$(B'C/A'C)^4$. If the antennas are both tilted an angle α , and $\alpha > \varphi_A$, then we have

$$\begin{aligned}\theta_A &= \alpha - \varphi_A \\ \theta_B &= \alpha - \varphi_B\end{aligned}\quad (5)$$

Since $\varphi_A > \varphi_B$, which leads to $\theta_A < \theta_B$ in (5),

we have not only $(B'C/A'C)^4 \gg 1$ but also

$[F(\theta_A)/F(\theta_B)]^2 > 1$ in (1). It essentially increases

the $\frac{C}{I}$ value. The increment of $[F(\theta_A)/F(\theta_B)]$

in (1) is decided by the slope of the vertical pattern factor, namely the derivative of $F(\theta) \sim \theta$ curve of

the antennas at $\theta_{ave} = (\theta_A + \theta_B)/2$, where

$\theta_A \approx \theta_B$. Generally, the nearer the θ_{ave} approaches the first zero of the pattern, the greater the slope of the curve is. Since φ_A and φ_B are

constant, the angle θ_{ave} can only be changed by the tilt angle α . This shows that α should be carefully adjusted for reducing the cochannel

interference to a permissible level at a specific area. In order to illustrate how to reduce cochannel interference between cells by the effect of tilted antennas, it is assumed that the cochannel cells where 120° directional antennas are used, are located at the shadow areas shown in Fig. 2.

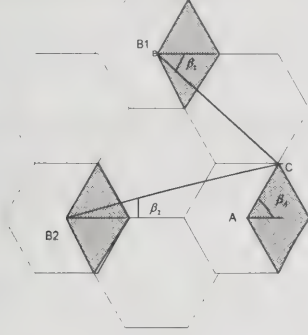


Fig.2 Cochannel interference between the cells.

A mobile unit C communicates with Station A and cochannel interference comes from Station B_i ($i=1,2$). It is also assumed that all the parameters of the antennas are the same as those stated above, and they emit the same ratio power. The carrier-interference ratio of the unit C to Station B_i can be derived as

where α_A' and α_{B_i}' are the effective tilt angles of Antenna A' at the direction β_A and of Antenna B_i' at the direction β_i respectively.

According to Ref. [2], we have

$$\alpha_A' = \cos^{-1} [1 - \cos^2 \beta_A (1 - \cos \alpha)]$$

$$\alpha_{B_i}' = \cos^{-1} [1 - \cos^2 \beta_i (1 - \cos \alpha)] \quad (8)$$

where α is actual tilt angle of the antennas. Similar to (3) and (4)

$$\varphi_A = \tan^{-1} \left(\frac{h_t}{AC} \right)$$

$$\varphi_{B_i} = \tan^{-1} \left(\frac{h_t}{B_i C} \right) \quad (9)$$

$$A'C = \frac{AC}{\cos \varphi_A}$$

$$B_i'C = \frac{B_i C}{\cos \varphi_{B_i}} \quad (10)$$

$$\left(\frac{C}{I} \right)_i = \left(\frac{B_i'C}{A'C} \right)^4 \left(\frac{F(\theta_A)}{F(\theta_{B_i})} \right)^2 \left[\sin \left(\frac{\pi}{2} \cos \beta_A \right) / \sin \left(\frac{\pi}{2} \cos \beta_i \right) \right]^2 \quad i = 1, 2 \quad (6)$$

where β_A and β_B are the azimuthal angles of the stations B_A and B_i respectively. Similar to (5), the angles θ_A and θ_{B_i} can respectively be expressed as [1]

$$\theta_A = \alpha_A' - \varphi_A$$

$$\theta_{B_i} = \alpha_{B_i}' - \varphi_{B_i} \quad (7)$$

The carrier-interference ratios $\left(\frac{C}{I} \right)_1$ and $\left(\frac{C}{I} \right)_2$ for different α can be calculated by (6) to (10). The total carrier-interference ratio can be calculated by

$$\left(\frac{C}{I} \right) = \left[\left(\frac{C}{I} \right)_1^{-1} + \left(\frac{C}{I} \right)_2^{-1} \right]^{-1} \quad (11)$$

As an example, assuming $AC=1000m$ in Fig. 2,

$h_1 = 50m$. We calculate geometrically that

$$B_1C = 2646m \quad B_2C = 3606m, \quad \text{and}$$

$$\beta_1 = 40.9^\circ, \quad \beta_2 = 13.9^\circ, \quad \left(\frac{C}{I}\right)_1 \text{ and } \left(\frac{C}{I}\right)_2 \text{ for}$$

different α can be calculated by (6) to (10), and tabulated in Table 1. The total carrier-interference

ratio at $\alpha = 22^\circ$ can be calculated by (11). The

result is $\left(\frac{C}{I}\right) = 17.59dB$, which meets the

demand of carrier-interference ratio in mobile communications.

Table 1 The results of calculation

α	0°	22°	25°
$\left(\frac{C}{I}\right)_1 (dB)$	14.40	18.09	19.28
$\left(\frac{C}{I}\right)_2 (dB)$	19.12	27.22	30.56

In practical application, the parameters of the neighboring cochannel base-station antennas are generally not the same, the tilt angle α of the main lobe of the antenna is very hard to be decided accurately by calculations. A work-site testing approach as the antenna being installed is proposed as follows. For example, in the above case (as shown in Fig.2), when Antenna B'_1 is installed, the field intensity emitted by Antenna B'_1 is tested at

the location C , while Station A and B_2 are turned off (not emitting). As increasing the tilt angle

α of the main lobe of Antenna B'_1 gradually, the field intensity at C tends weaker. But the tilt should not proceed before the minimum of intensity reading occurs. Otherwise the main lobe of Antenna

B'_1 will be over tilted and the first side lobe will take effect with the result that interference at the location C will increase and a blind area in the service area of Station B_1 will appear. Thereafter

Station B_1 is turned off, and Antenna B'_2 is

adjusted in the same procedure as Antenna B'_1 . For

best cochannel interference reduction, the base-station antennas should have some device for fine-adjusting of their tilt, and some communications should be established between the work-site of antenna installation and the field intensity testing spot.

Conclusion

The cause of a tilted base-station antenna being of benefit to cochannel interference reduction is studied analytically in this paper. The procedure of adjusting the tilt of the antenna is also theoretically explained. It is expected that a proper adjusted transmitting antenna of base-station will play a more important roll in cochannel interference reduction in mobile communications.

REFERENCES

- [1] W. L. Stutzman, *et al.*, *Antenna Theory and Design*, John Wiley(1981)86-125
- [2] W. C. Y. Lee, *Mobile Cellular Telecommunications*, McGraw-Hill(1995)205-208

Electronic Beam Tilting of Base Station Antenna : Rotman Lens Fed

S. G. Kim, S. M. Park, C. W. Ro, and J. M. Lee

Gamma Nu Wave, Inc.

395 Woncheon-Dong, Paldal-Gu, Suwon-City, Kyungki-Do, Korea, 442-380

82-331-213-9217, 82-331-213-9219 (fax), gammanu@chollian.net

Abstract

A new electronically controlled down-tilt mechanism for wireless communication base station antenna is presented. It's main feature consists of a microwave switch, a symmetrical bootlace lens which is a beam forming device like Rotman lens, and linearly arrayed antennas. Each beam port of the symmetrical bootlace lens is designed to give a linear phase front across the radiating elements, yielding a radiation spatially summed in a direction normal to the phase-front. Furthermore, a few watts small power amplifiers may be inserted between the array ports of the lens and radiating elements to replace the conventional single high power amplifier and heavy cable runs and fixedly directive sector antenna with a cost-effective tower-top phased array antenna. These concepts are checked by experimental results.

Key words: Base station antenna, Electronic Down-tilt, Beam forming, Rotman lens, Phased array

Introduction

Even though the base station antenna amounts to only a few percent of the overall cost of a wireless communication site such as cellular, PCS and wireless local loop, etc, its performance effect is enormous since it determines directly the air interface quality to and from the remote users. There are many concerns about the base station antenna; gain, polarization, null-fill, upper sidelobe suppression, and down-tilt.

In those concerns, tilting the main beam below the horizon has often been requested to reduce the interferences between two frequency reuse cells and to improve the system functions in view of multi-path distortion. However, the conventional method using the mechanical down-tilt bracket as in Figure 1, has been one of the cumbersome burdens in the cell site management as well as in the installation. According to [1], it can cost upwards of \$2,000 to rent a crane and \$3,000 a day for a crew of tower climbers and riggers to go to an antenna site and adjust the mechanical bracket. What is worse, the site may have to be taken off or

traffic rerouted during this process. Obviously, it is expensive, even dangerous, degrades the service quality resulting in customer dissatisfaction.

In this paper, a new base station antenna is introduced as a means to replace the conventional mechanical down-tilt with electronic one controllable at the remote cell site far from the base station. Its main feature consists of a microwave switch and a microwave lens, and a linear array antenna.

Beam Forming Theory

Microwave lens introduced by Rotman and Turner[2] or symmetrical boot-lace lens by Shelton[3] may be used to feed a linear array antenna to produce multiple beams. It resembled an optical lens offering collimation of the light emitted from the focal point. However, microwave lens is multi-focal, different from single focus optical lens. It is the key to produce multiple down-tilt beams.

A microwave lens consists of a 2-dimensional parallel plate waveguide and a set of transmission lines through which TEM

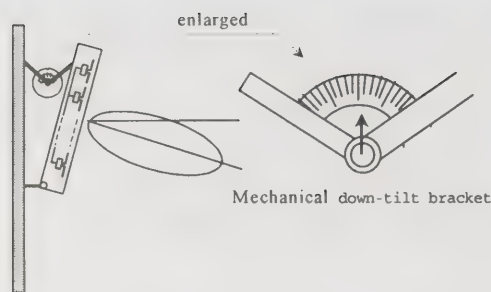


Fig.1 Conventional mechanical down-tilt base station antenna

mode can propagate. On one side of the parallel plate structure, there are a few beam ports and on the other side a few array ports which feed a linear array antenna as shown in Figure 2. The coordinates and contours of the parallel plate waveguide and the lengths of the boot-lace transmission lines may be designed so that the optical path lengths from each beam port will give a linear phase-front across linearly arrayed antennas. It steers the beam into the direction normal to the phase-front.

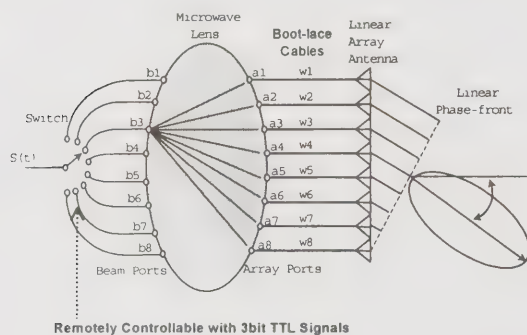


Fig. 2 Lens fed electronic down-tilt base station antenna

Whenever down-tilt change is needed, a digital command from the remote cell site may be sent to the base station via wireless telephone, etc. It is decoded as switch control signals, like TTL, by the software driven tilt manager and runs to the antenna tower by signal cables to enable the amplified RF signal coming from the base station to go through the selected beam port producing a steered beam in the desired direction.

Lens Design

A symmetrically configured boot-lace lens[3] was designed to operate at 2.35GHz wireless local loop band. It has microstrip parallel plate structure and boot-lace transmission lines etched on a high dielectric constant low loss laminate RO3010 ($\epsilon_r = 10$), offering cost-effective small size, referenced with a pack of long cigarette in Figure 3.

In the total 14 beam ports, 8 center beam ports provide 8 fine steered beams about 2.8° spaced to cover $\pm 10^\circ$ angular range from the horizon. The other both side 6 ports are dummy ports extensible up to $\pm 15^\circ$ coverage. According to the guideline[4], maximum down-tilt was limited up to 10° to prevent the beam in the boresight direction from being distorted too much. Obviously, it is narrow angle fine resolution requirements formidable by any other phased array techniques, for example, such as by active phase shifter control.

As in Figure 3(c), only 8 center array ports of all 14 ones were used to feed 8 antenna elements by equal length cables. The antennas are linearly arrayed in half wave lengths spacings at the highest frequency in the operating band to avoid grating lobes. Here, aperture coupled microstrip antenna[4] were developed as antenna element for high gain and broadband radiation. Both side 6 dummy ports may be used to extend the number of array elements up to 14 in case higher effective radiated power is required.

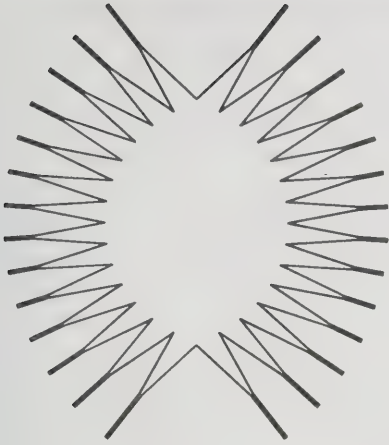
A vast design and simulation processes for beam forming microwave lens for digital cellular, PCS, and FPLMTS as well as for wireless local loop band were established by this study. When the narrow angle fine resolving down-tilt is required, as in this case, the symmetrical boot-lace lens is more attractive to be realized in a cost-effective small size than the conventional Rotman lens. In case of Rotman lens, the microstrip beam ports are too close each other to be impedance matched in a small size than in the symmetric boot-lace lens. Due to this reason, the Rotman lens becomes larger than the symmetrical boot-lace lens.

Experiments

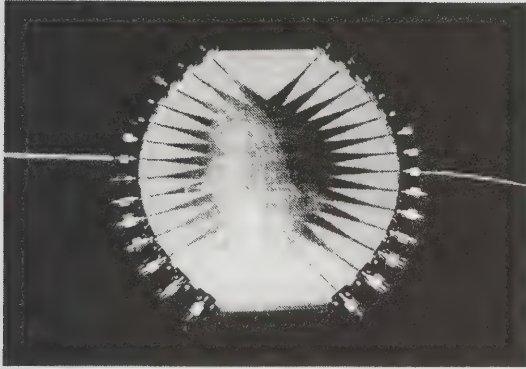
1) Lens' Mag/Phase Transfer

The network transfer characteristics through the lens is important since it gives uniform excitation and aberration free beam steering. In Figure 4, the magnitude and phase transfer from a beam port to 8 array ports were measured by the calibrated vector network analyzer and overlapped with those from the other beam ports for comparison.

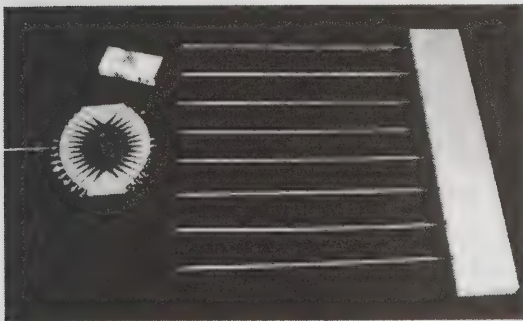
Magnitude transfers from beam ports to array ports are shown to be within $\pm 1\text{dB}$ deviation so that the lens will excite all the antenna elements in uniform magnitudes. Phase transfers from a beam port to array ports shows almost linear phase front, as expected in geometric optics path length differences, producing almost aberration-free fine resolution beam steering.



(a) Symmetrical boot-lace lens artwork

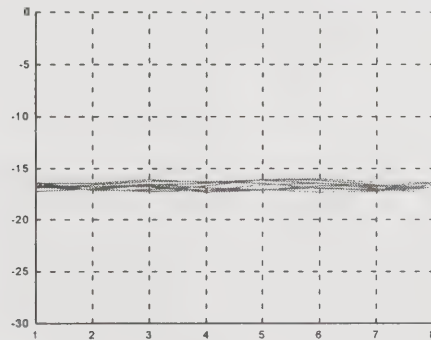


(b) Lens realized in microstrip line type using $\epsilon_r=10$ laminate and microstrip connectors

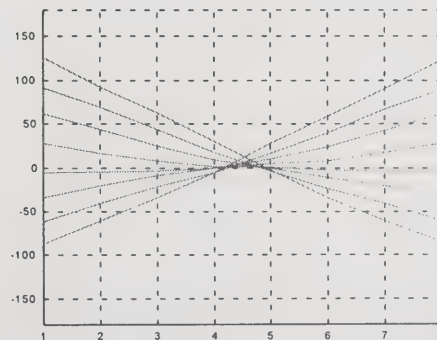


(c) Configuration feeding 8 linearly arrayed aperture coupled microstrip antennas

Fig.3 Lens fed down-tilt antenna at 2.35 GHz wireless local loop bands.



(a) Magnitude transfer



(b) Phase transfer

Fig.4 Overlapped Mag/Phase transfers of designed lens : S21 measured from beam ports to array ports.

2) Beam steered radiation patterns

The beam steered radiation patterns were measured in a far-field anechoic chamber as in Figure 5. It shows explicitly almost equally spaced 8 steered beams covering $\pm 10^\circ$ elevation angles. Clear fine beam steering and sidelobe levels around -13dB below the main steered beams state that 8 linearly arrayed antennas were well excited by the designed lens with uniform magnitudes and linear phases as expected.

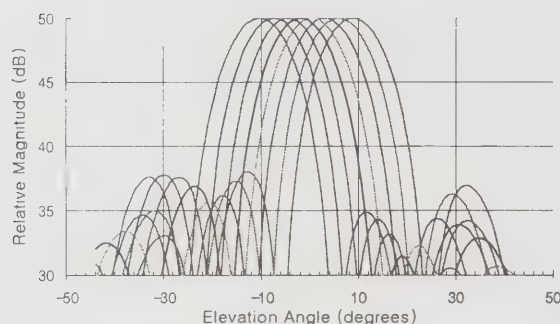


Fig. 5 Beam steered far-field radiation patterns : measured.

3) Matching and insertion loss of the lens

However, this study remained a problem. Typical VSWR higher than 3:1 was observed as in Figure 6. This is due to severe mutual coupling between microstrip ports. The gap between ports is too close, less than a quarter wave length, to be well matched to the 50Ω lines. It may be inevitable in these kind small size narrow angle fine resolving requirements. However, a transformer or tradeoff with size allowing good match should be devised.

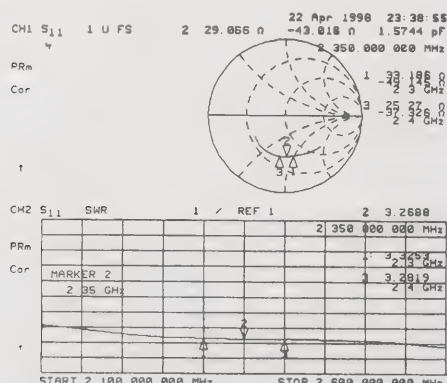


Fig.6 Input impedance of the lens

Conclusion

By adopting a microwave lens and a microwave switch, a new phased array base station antenna was proposed to replace the conventional mechanical down-tilt with electronic and remote controllable one so that the traffic can not be taken off or re-routed whenever the down-tilt change is desired. It would reduce the tilt control complexity and cost in the cell site maintenance as well as in the installation.

The proposed method provides narrow angle fine resolution beam steering which is cost-effective than that by any other phased array technologies, for example, by active phase shifter control.

This study may be extended to include an array of small watts solid state power amplifiers between array ports of the lens and the linearly arrayed radiating elements, providing an in-expensive tower-top active phased array antenna system. It may replace expensive conventional single high power or linear power amplifier and high directive sector antenna in cellular, PCS, and wireless local loop system.

References

- [1] R. Wilson, "Selecting Antennas for a Wireless System," *Wireless Design and Development*, pp. 81-86, Feb 1988.
- [2] W. Rotman and R. Turner, "Wide-angle microwave lens for line source applications," *IEEE Trans. Antennas and Propagat.*, Vol. AP-11, No. 6, pp. 623-632, Nov 1963.
- [3] J. P. Shelton, "Focusing characteristics of symmetrically configured bootlace lenses," *IEEE Trans. Antennas and Propagat.*, Vol. AP-26, No. 4, pp. 513-518, July 1978.
- [4] D. M. Pozar, "A Reciprocity Method of Analysis for Printed Slot and Slot-Coupled Microstrip Antennas," *IEEE Trans. Antennas and Propagat.*, Vol. AP-34, No. 12, pp. 1439-1446, Dec. 1986.
- [5] A. Singer, "Selecting antennas for PCS and DCS systems," *Mobile Radio Technology*, pp. 10-22, July 1997.

An Improved 3D Ray Tracing Method for Indoor Propagation Prediction

W.Lu , K.T.Chan

Department of Electronic Engineering, The Chinese University of Hong Kong
Shatin, NT, Hong Kong

(852)26098257, (852)26035558(fax), wlu@ee.cuhk.edu.hk

Abstract

The 3-D brute-force ray-tracing technique has been used extensively to predict indoor/outdoor radio propagation in mobile communication in recent year. This method produces accurate, deterministic channel models for wireless system design. It is also a powerful tool for propagation prediction in complex environments. However, a reception sphere has to be introduced to determine the rays that reach the receivers and the kinematic errors unavoidably occur. This paper presents a novel shooting and reception sphere technique which provides a much improved accuracy.

Key words: Ray tracing, Radiowave propagation

Introduction

The characterisation of propagation conditions for mobile radio systems design based on many models^[1-5], has been widely studied with the emergency of a variety of wireless indoor system. For "high cost and high density" personal communication networks, in office buildings for example, 3D site-specific ray tracing propagation models would be of great interest. Many propagation experiments during the past few years have proved that the 3D ray tracing models can provide unpresented accuracy.

However, one difficulty of this method is how to find a reception sphere with correct radius which can receive exactly one ray each time. If the radius is too large or small, kinematics errors due to double count or zero count would occur. This paper advance an improved ray launching and novel reception sphere technique. The kinematics errors in this method are reduced significantly.

3D Ray Tracing

A. Ray Launching

The conventional solution for the source ray directions is adapt from the theory of geodesic dome. Each triangular face of the icosahedron which is inscribed inside the unit sphere is tessellated into N equal segments where N is tessellation frequency. All of rays are launched at angles that pass through the vertices of the triangles. In spherical co-ordinate system, we can easily define a direction vector $\vec{R} = \begin{bmatrix} \phi \\ \theta \end{bmatrix}$, in

which ϕ resembles a meridian of longitude and θ resembles a specification of latitude. For a specific triangular face of icosahedron, each vertices location can be determined by the method of [6] as

$$\vec{R} = \begin{bmatrix} \phi \\ \theta \end{bmatrix} = \begin{bmatrix} \arctan(x_1 / y_1) \\ \arctan \sqrt{x_1^2 + y_1^2} / z_1 \end{bmatrix} \quad (1)$$

In our improved ray tracing method, because of the modification of reception sphere technique, we also make some justification on the ray launching method. Besides those regular rays, the rays that pass through the centers of tessellated triangles, shown in Fig.1a, are also included. In the calculation of the ray launching vector, we should increase the tessellation frequency N to $3N$. Fig1b provides an example in regular tessellation frequency 2. Only the direction vectors of those black dots are what we should calculated.

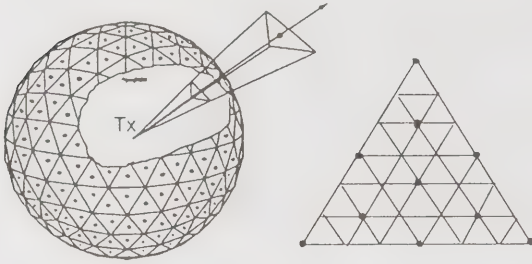


Fig1 a) Ray launching from geodesic vertices

b) A tessellated triangle

B. Reception sphere

The regular method for interpreting the traced ray information is the reception sphere model. Rays that intersect one sphere contribute to the total received power. To guarantee the collection of at least one ray from the source, the minimum radius for a reception sphere is $1/\sqrt{3}$ the distance between rays. However, this radius sweeps out a circular area across the wavefront where sometimes two rays intersect the sphere. So, double count error which register additional field at the particular receiver location occurs with a probability of 20.9%.^[7]

Instead of using the radius in regular reception sphere technique, a smaller radius of reception sphere which is only $1/2$ the distance between adjacent regular rays is applied here. These spheres are tangential with each other. To identify the reception of the rays which pass through the centers of triangular faces, the tiny spheres are introduced with a radius of $\sqrt{3}/3 - 1/2$ times the distance between adjacent regular rays, as shown in Fig.2.

Obviously, the double count errors would not occur. But the zero count error will occur with a probability P , which is given by Equ.2:

$$P = \frac{S_z}{S_\Delta} \approx 4.97\% \quad (2)$$

Where $S_\Delta = \sqrt{3}/4a^2$

$$S_z = S_\Delta - \pi a^2 (17/24 - \sqrt{3}/3)$$

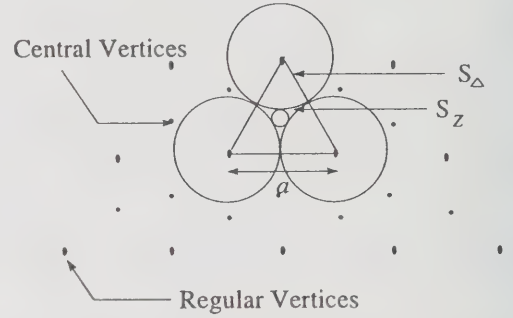


Fig2. Modified reception sphere zero counting(S_z/S_Δ)

Compared with that of the double count error, the probability of zero count is reduced significantly.

Results

This modified ray launching and reception sphere technique was tested in a simple theoretical environment. In Fig3, we consider a furnished room which is 3-m high, 21-m long and 12-m wide. Because metallic furniture is the most dominant reflector in furnish room, we place a metal closet with 2-m high in this room. The emitting antenna is located at 0.5m from ceiling while the mobile receiving antenna follows a predetermined path at a height of 1.5m. Table I give the realistic values of dielectric constant we use in our simulation. The received signal power of the receiver line is given as:

$$P_r = P_t \frac{\lambda^2}{(4\pi)^2} \left[\frac{1}{r_{LOS}} + \sum_i \frac{\exp(j2\pi r_i / \lambda)}{r_i} R \right]^2$$

Where $P_{r,i}$ = The received power or transmitted power
 r_{LOS} = The LOS distance between source and receiver
 r_i = The i th reflection path length

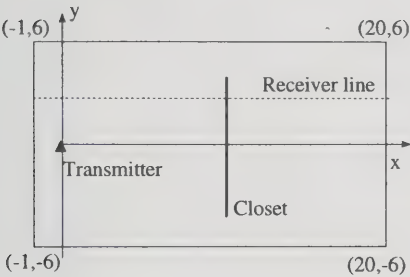


Fig3. A simple testing rectangular room with a metal furniture.

Material	ϵ_r
Walls(rear and front) (gyproc)	8
Walls (right and left) (stucco)	36
floor(linoleum)	10.5
ceiling (metallic tiles)	

Table I. Dielectric constant of the room.

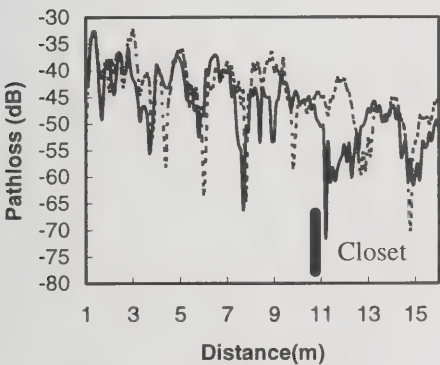


Fig4. Power along receiver line using regular and modified reception sphere

----- Regular method
 _____ Modified method

The received power after the closet should be lower than that in the empty room as some specula ray paths which exist in the empty rooms are blocked by the furniture. It is observed from Fig.4 that the regular reception

sphere method results in a higher power level behind the metal furniture due to double count errors while the modified method provides a more realistic power level.

Conclusions

The use of this modified geodesic sphere methods increase the accuracy of 3D ray tracing for propagation prediction. The simulation results can correctly predict the signal field distribution trend which is useful for radio system designers to optimise coverage minimise interference and optimise frequency re-use in an indoor environment. The model based on modified geodesic sphere methods offer a reliable solution to the study and simulation of propagation characteristics in indoor environments.

Reference:

- [1] R.E. Collin, *Field Theory of Guided waves*, New York: Mcgraw-Hill,1960,ch.1.
- [2] M.A.K.Hamid and W.A.Johnson, "Ray-optical solution for the dyadic Green's function in a rectangular cavity," *Electron Lett.*, Vol.6, no.10., pp.317-319, May 1970
- [3] Walter Honcharenko, H.L. Bertoni: 'Mechanisms Governing UHF Propagation on Single Floors in Modern Office Buildings'. *IEEE Trans.Veh.Technol.*, 1992,**41**,(4), pp.496-503
- [4] S.Y.Seidel and T.S.Rappaport, "A ray tracing technique to predict path loss and delay spread inside room," in *Proc. IEEE GLOBECOM Conf.*, Orlando, FL, Dec.,1992, pp.649-653
- [5] S.H.Chen and S.K.Jeng, "An SBR/Image approach for indoor radio propagation in a corridor," in *Proc.Asia-Pacific Microwave Conf.*, Tokyo, Japan, Dec.1994, pp.205-208
- [6] H.Kenner, *Geodesic Math and How to Use it*. Berkeley, CA, Univ. of California Press, 1976
- [7] G.Durgin, N.Patwari: 'Improved 3D ray launching method for wireless propagation prediction', *Electron.Lett.*, 1997, **33**, (16), pp1412-1413

Miniature Dielectric-loaded Personal Telephone Antennas with Low SAR

Bernhard Rosenberger, Rosenberger HF - Technik GmbH & Co., P: + 49 8684 - 180, F: + 49 8684 - 18197
Email:BERND@Rosenberger.de

1. Introduction

Over the short history of the cellular telephone we have witnessed astonishing advances to the state of the art of personal telephone handset design. However the antenna solutions that are currently employed are usually rather traditional structures and as the size and weight of handsets have reduced the awkward realities and limitations of these structures have become noticeably more troublesome to designers. These difficulties are likely to increase as the public anxiety about possible health risks of using personal telephones is mounting and it is becoming increasingly important to produce handsets which completely meet the international S.A.R. (Specific energy-Absorption Rate) regulations.

The most common antenna topologies used on handsets today are basically monopole or monopole derived structures. Actually clean classification of hand-set antennas is complicated by the fact that in reality hand-set cases are far too small to present sufficient ground-plane mirroring for a classical monopole and often forms of "mirror" element loading are employed to improve performance. The imprecision by which personal telephone handsets are rated for antenna performance is an artefact for this historical use of the monopole. In reality the handset body is actually functioning as part of the antenna and the free-space gain of the antenna can really only be properly measured with the antenna attached to the handset. Of course the validity of the characterisation is placed in doubt by the presence of the user's hand grasping the handset as well not to mention

the effect of the user's head. These difficulties have been overcome from a type-rating point of view by requiring the handset to comply with specified power into 50 ohm load at the antenna connection port. This neatly avoids the need to regulate the effective EIRP of the handset which might be more indicative of its communication performance.

One significant limitation of monopole and some monopole derived antennas concerns their lack of polarisation diversity. Personal telephone handsets are often used inside buildings and in urban environments where scattering of the signal from vertically and horizontally disposed surfaces promote mixing of polarisations which changes spatially. Moreover the orientation of the handset with respect to the dominant polarisation plane of the base-station antenna can be entirely arbitrary. Efforts¹ have been made to propose "figure of merit" measures for mobile telephone antennas in cluttered urban environments. This work has shown that antennas with high polarisation diversity have superior practical communications performance.

The Twisted Loop Antenna

The twisted loop or bifilar helix antenna is designed to minimise the S.A.R. It is basically a loop antenna which is balanced at the top by a co-axial feed which passes axially through the centre. The lower portion of the antenna is a sleeve balun which creates a transition between single-ended and balanced drive so that the loop antenna can be fed with balanced currents. Currents at the centre conductor feed point of the antenna pass round a 360° path including the rim of the sleeve balun and then back to the outer conductor of the co-axial cable feed.

In a number of respects this topology is, without dielectric loading, a low S.A.R. structure. As can be seen from figure 1 the twisted loop antenna creates a voltage maximum half-way up each helix for the balanced mode with the net effect that an equivalent voltage dipole is resolved across the diameter of the antenna in a direction which meets the two helices at their voltage maxima. For optimum S.A.R. performance the antenna should be orientated such that this dipole is placed normal to the surface of the face. In this direction the antenna projects an electrostatic minimum into the face and also the far-field radiation pattern in this direction is also a minimum. The 180° twist in the twisted loop topology provides the characteristic that the axis of electrostatic field minimum and also far-field radiation pattern minimum are the same direction whereas for the conventional planar loop these two axes are orthogonal.

2. Standards and Testprocedures [6]

On the first of August 1996, the US Federal Communications Commission (FCC) issued a report and order [1] which requires routine SAR evaluation of mobile telecommunications devices prior to equipment authorisation or use. In Europe, a working group of CENELEC mandated by DGXIII [2] will be submitting their report and recommendations on the technical requirements for compliance testing of handheld mobile telecommunications equipment (MTE) [3]

in the next few months. The International Commission on Non-Ionizing Radiation Protection recommends that MTE devices should meet the spatial peak SAR limits set by ICNIRP [4]. In view of these developments there can be no doubt of the pressing for the industry to establish a reliable method for assessing the spatial

peak SAR values induced by MTE in users.

Various dosimetric techniques have been developed to perform such compliance testing [5]. However, none of these procedures have been standardised, due to serious shortcomings in the development of the methodology needed to assess inherent overall errors. A thorough assessment of the uncertainty is of critical importance, since the exposures due to current devices have been shown to be very close to and even above the stipulated safety limits. This problem is aggravated by the fact that exposure significantly depends upon the way the device is held and also to a large extent on the anatomy of the user. It is important from the point of view of manufactures and service providers that the adopted testing procedure does not drastically underestimate the maximum real-world exposure. This could otherwise lead to a loss of credibility in the eyes of the public, as well as to the prospect of damaging liability suits. On the other hand, the need for thorough testing should not impede development of new technologies or cause undue delay in marketing of devices which such short product cycle times as MTE.

3. Definition of Mobile Telecom Equipment (MTE) [7]

MTE are all types of telecommunications equipment classified as „mobile“ in the Radio Regulations of the ITU, Chapter 1 „terminology“, Article 1 „Terms and definitions“, Section IV „radio stations and systems“.

The list of the equipment falling under this category is comprised of the following categories: survival craft station), mobile station, mobile earth station, land mobile station, land mobile earth station ship station, ship earth station, on-board communication station, aircraft station,

aircraft earth station, radionavigation mobile station, radiolocation mobile station. Although this list assumes a fairly wide interpretation of the word „mobile“, the material herein concentrates solely on handheld portable wireless MTE.

4. Definition of Specific absorption rate (SAR). [7]

The time derivate of the incremental energy (dW) absorbed by (dissipated in) an incremental mass (dm) contained in a volume element (dV) of given mass density (p)

$$SAR = \frac{d}{dt} \left(\frac{dW}{dm} \right) = \frac{d}{dt} \left(\frac{dW}{p dV} \right) \quad (4.1)$$

SAR is expressed in units of watts per kilogram (W/kg).

Note: SAR can be calculated by:

$$SAR = \frac{\sigma E_i^2}{p} \quad (4.2)$$

$$SAR = c_i \frac{dT}{dt} \quad (4.3)$$

$$SAR = \frac{j^2}{p\sigma} \quad (4.4)$$

where

E_i : rms value of the electric strength in the tissue in V/m

σ : conductivity of body tissue in S/m

p : density of body tissue in kg/m³

c_i : heat capacity of body tissue in J/kg K

$\frac{dT}{dt}$: time derivative of temperature in body tissue in K/s

J : magnitude of the induced current density in the body tissue in A/m²

5. Basic Requirements [6]

There has been widespread discussion in the industry, health agencies and standards committees as to whether compliance testing must be conducting considering worst-case conditions. The pros and cons of this discussion have been summarised in [5].

However, quite apart from the definition of the operational conditions under which compliance must be achieved, the applied technique and procedure should also satisfy the following basic requirements:

- ensure the assessed SRA values do not underestimate the exposure of a reasonable cross section of the MTE user group including children;
- provide reproducible results independent of the testing laboratory;
- allow for testing the effects of manufacturing variabilities on the SAR values; and
- be unbiased with respect to the design of MTE, i.e., MTE devices resulting in low SAR values in the tests should actually be reflected by reduced exposure in the real-world situation and vice-versa.

5.1. Preliminary Considerations [7]

This section describes the measurement simulation and modelling requirements for the procedures involved in testing compliance of MTE with the basic restrictions.

Measurement and computational simulation procedures should be considered as suitable approaches . In case of numerical compliance testing, at least one position of mobile phone should be verified by measurements. The testing of compliance with the basic restrictions requires special considerations, particularly in extremely close near fields.

Compliance with the basic restrictions can be done directly determining the spatial distribution of the specific absorption rate (SAR) induced in experimental or theoretical phantoms representing the human body, from which the maximum SAR averaged over any 10 g of tissue shall be determined.

Procedures described in this section shall be used to demonstrate compliance with the basic restrictions.

5.2. Practical Testing [6], [7]

The method should ensure that the assessed SAR values do not underestimate the exposure of a reasonable cross section of mobile telecommunications users, including children, under the tested operational conditions. It should also take into account the manufacturing tolerances. The MTE should be tested under the operational condition at the highest power at the central transmitting band frequency. Tests shall be performed for each configuration (e.g., with all antennas provided) and for each system with the MTE can operate (multi-system devices).

Test procedures should give reproducible results for all devices and various operating conditions.

5.3. Operational Conditions for Certain classes of Devices

The technical realisation and handling differ significantly for types of MTEs. Hence, the operational conditions for each class of MTE must be well defined for testing.

Handheld Mobile Phones

In this context handheld mobile phones are those cellular telephones consisting of one unit with integrated microphone, ear piece, transceiver and antenna. When testing the

compliance of these devices with basic SAR values, measurements must be taken under the exposure conditions specified below.

Intended Use

The intended or normal operating condition of the MTE is generally established by fully extending the antenna of the MTE, i.e., when the case plus the antenna are extended to the maximum possible overall length.

The intended use or normal position is a position which is convenient and provides good acoustic coupling. If the manufacturer does not specify the intended use position, then it is defined as follows :

- The centre of the ear piece shall be placed directly at the entrance of the auditory canal.
- The reference line of the phone is defined to be the „tangential“ line (on the surface of the phone's case facing the phantom) which connects the centre of the ear piece with the centre of the bottom of the case (typically near the microphone). Exceptions would be foldable phones or phones with flip covers, for which the reference is defined to be the ear piece with the centre of the hinge. For strongly curved phones („banana“ shaped phone), the reference line must be specified by the manufacturer.
- The reference line defined above shall lie in the reference plane defined by the following three points: auditory canal openings of both ears and the centre of the closed mouth. The plane of the ear piece of the devices is normal to the reference plan.

5.4. Test Operational Conditions and Positions

Since it is often difficult for the user of a handheld mobile phone to precisely control the position of the device with respect to the head, considerable variations of this intended use cannot be avoided. Parallel shifts as well as different orientations must be expected. Since the range of parallel shifts can be covered by the requirements of the modelling of the head, the following additional operational conditions shall cover the variation range of the intended use operation:

- The intended use position is defined by an angle of 80° between the reference line of the phone and the line connecting both auditory canal openings. If this position is not applicable, then it may be deduced by referring to the user's manual.
- The angle between the reference line of the phone and the line connecting both auditory canal openings shall be reduced until the device touches the face of the phantom.
- The angle between the reference line of the phone and the line connecting both auditory canal openings shall be increased to 100° , provided that the antenna does not touch the head. If the antenna does touch the head before an angle of 100° has been reached, this smaller angle shall be applied.
- The reference plane shall be tilted by 30° in the direction of the body's axis, whereby the angle between the reference line of the phone and the line connecting both auditory canal openings shall be 90° .

6. The measurement of SAR for different frequency ranges on dielectric loaded twisted loop Antennas

The objective of the measurements was to study the suitability of the twisted-loop antenna type for use in handheld mobile communications equipment with respect to radiation exposure of the user. The study uses the following near-field measurements of the antennas mounted on a generic phone:

- Measurement of the magnetic near-field in free space and in the flat phantom with tissue simulating liquid.
- SAR measurement in the flat phantom
- Assessment of the spatial peak SAR - values in the head phantom according to the WGMTE CENELEC TC211B : final draft, Feb. 1997 document and the FCC preliminary standards described in the IEEE SCC34 documents.

All the measurements were performed at the ETH in Zurich with the DASY3 dosimetric assessment system and near-field scanner. The magnetic fields were measured with the isotropic H-field probe H3DV4; SAR were measured with the dosimetric probe ETDV5. Both probes were calibrated and system validation checks were performed prior to measurements. Phantom-based measurements were performed in a new Generic Twin Phantom with 2 mm shell thickness.

The test specimens were Dielectric loaded Twisted-loop Antennas for GSM and DECT (Digital European Cordless Telephone) at their specific frequencies:

- GSM 11 : 906 MHz
- GSM 14 : 908 MHz
- DECT 14 : 1750 MHz

The antenna input power was set to 250 mW unless otherwise indicated in the measurements. (forward or available power at the antenna connector).

6.1. Testresults :

The measured results prove the low SAR characteristics of this type of antenna .

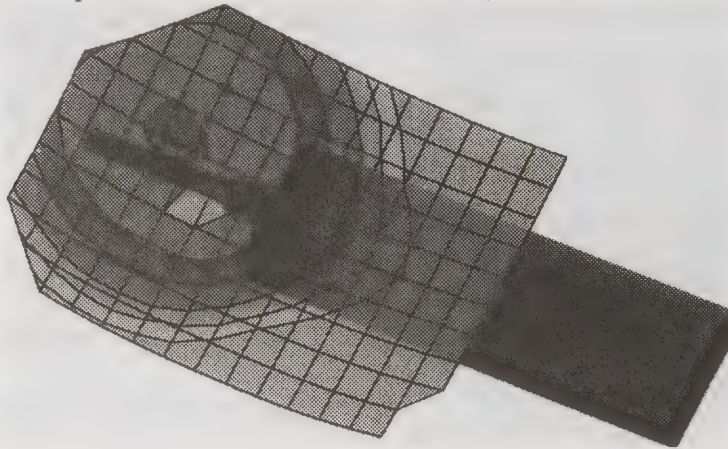


Fig 1 SAR measurement at the surface of the dummy head.

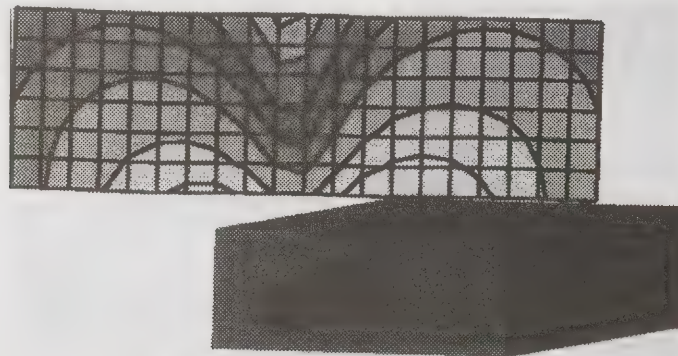


Fig 2: Field distributes normal to the users head .

References:

- [1] Federal Communications Commission. "Report and order: Guidelines for evaluating the environmental effects of radiofrequency radiation", Tech. Rep. FCC 96-326, FCC, Washington, D.C. 20554, 1996.
- [2] Directorate-General III (Industry), Commission of the EC, Standardisation mandate forwarded to CEN/CENELEC/-ETSI in the field of information technology and telecommunications, EU, 1996.
- [3] WGMTE of CENELEC TC211/B, Final Draft: Considerations for human exposure to electromagnetic fields from mobile telecommunication equipment (MTE) in the frequency range 30 MHz-6 GHz, CENELEC, Brussels, Feb. 1997.
- [4] International Commission on Non-Ionizing Radiation Protection (ICNIRP), "Health issues related to the use of hand-held radiotelephones and base transmitters", *Health Physics*, vol. 70, no.4, pp. 587-593, Apr. 1996.
- [5] N. Kuster, Q. Balzano and J. I. Lin, eds., *Mobile Communications Safety*, Chapman & Hall, London, 1997, p.289.
- [6] N. Kuster, R. Kästle and Th. Schmid: Dosimetric Evaluation of Handheld Mobile Communications Equipment with Known Precision, Swiss Federal Institute of Technology, IEICE Trans., Vol. E80-B, No. 5 May 1997.
- [7] Final Draft: Considerations for evaluation of human exposure to Electromagnetic Fields (EMFs) from Mobile Telecommunication Equipment (MTE) in the frequency range 30 MHz-6 GHz, CENELEC, Brussels, Dez.1997.

The Author :

Bernhard Rosenberger , 44 , is head of the research and development at Rosenberger Hochfrequenztechnik , Germany . His background is an engineering degree in microwave electronics from Munich university. After graduation he has worked on the development and standardization off for coaxial connectors in frequency ranges up to 110 GHz. He headed several German research programs for microwave multi-chip module packaging and high speed transmission systems. He and his company contribute to the IEC TC for passive intermodulation.

A Practical Approach to Wideband Antenna Array Characterization

S. London, V. Leonov, P. Koert, L. Susman

Advanced Power Technologies Inc.,

1250 24th Street NW, Washington DC 20037

Tel: (202) 223-8808, fax: (202) 223-1377, e-mail: simon@apti.com, vleonov@apti.com

Abstract

A high power HF phased antenna array capable of very wideband frequency and scan coverage is being constructed as part of the High Frequency Active Auroral Research Program (HAARP), in Galkona, AK. This paper presents measurement and analytic array characterization approach for the purpose of antenna evaluation and antenna matching unit (AMU) design. It is based on the use of scattering parameters measurement of the array and has resulted in a 30% improvement in low frequency range of the HAARP system. The effectiveness of the design is further evaluated by calculating the average efficiency of the array system, which is within 1.5 dB of its theoretical maximum.

Key words: Antenna array, odd mode, even mode, hybrid, S matrix de-embedding.

Introduction

The HAARP Developmental Prototype array (DP) consists of a 6 x 8 array of lowband and highband cross dipole elements (the highband element being mounted under the corresponding lowband element), a total of 96 cross dipoles. The low and high band elements cover the frequency range, 2.8 to 7 MHz and 7 to 10 MHz, respectively and are capable of scanning up to $\pm 30^\circ$. [1]. A separate 10 kW transmitter energizes each dipole. DP array was built as a proof of engineering principle before the array expansion to its final 12 X 15 size is attempted. A remarkable feature of the array is its ability to function over the large range of antenna impedance variation encountered over the frequency/scan volume, attributed mainly to the AMU.

Extensive testing of a DP version of the array discovered undesirable feed resonance effects at discrete frequencies. Analyses and subsequent measurements confirmed even mode excitations of the dipole, generated by electrical asymmetries in the multiple elements, as the root cause of these difficulties. A new feed system design has been developed to solve this problem. The initial DP design had been obtained by NEC antenna simulation. Because of limited accuracy of the simulation a new array mathematical

characterization had to be accomplished based on measurements of the constructed antenna. Prior analysis and measurements had shown that a 5x5 array would be needed to represent a larger array. Under these circumstances a straightforward approach would be the measurements of the S-matrix between dipole's halves (to capture both the even and odd mode characterization of the array); this yields a 100 x 100 element matrix and results in numerical accuracy loss in the design process. To solve this problem an approach that separates odd and even mode excitation components has been developed and used. Specially, exciting the dipole through 180° hybrids provides an independent even mode termination while permitting measurements of the desired odd mode. This allowed a measurement program for the 50 x 50 odd mode matrix to be carried out in a practical time scale and with manageable accuracy. Even port termination has been experimentally optimized for resonance suppression and acceptably low power losses. For the full array characterization additional measurements could be done for the even mode ports and cross odd/even mode ports. A full S-matrix would allow feed system optimization for a specified level of even mode excitation. For the estimation of the even mode power only it is necessary to measure even and odd/even port couplings for a small part of the

array where even mode excitation is expected to be the significant.

Characterization Procedure

Characterization is done in the following steps:

- Measurement of the array odd mode S_m matrix including hybrids
- Extracting of antenna odd mode S_a matrix

The measurement block diagram is presented in Figure 1.

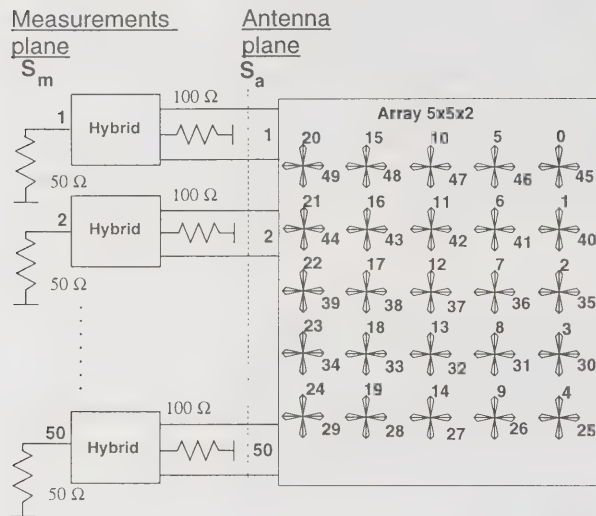


Figure 1. Antenna Measurement Block Diagram

Identical wideband transmission line hybrids were developed and installed directly at the dipoles. Hybrids were built according schematic shown in Figure 2 using 50 Ohms coaxial cable wound on ferrite cores.

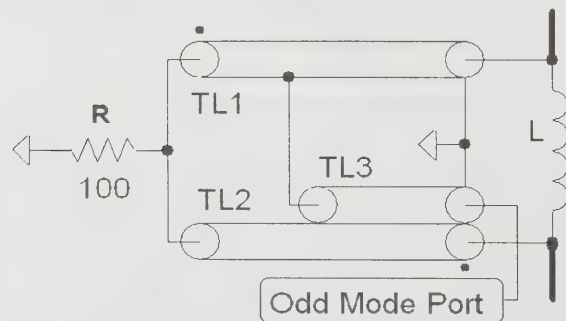


Figure 2. Measurement Hybrid

Inductance L and the TL3 tap point have been optimized for minimum odd mode port reflection (less than 0.6) in order to reduce errors in measurements and data processing. The value of the even mode termination resistor R has been

experimentally chosen to be 100 Ohms. Measurements of the odd mode S_m matrix were made using a HP Automatic Network Analyzer for 2-12 MHz band for both low and high band dipoles. The ANA has been calibrated up to the hybrid odd mode ports. The measurements resulted in 85 files. For quality insurance all coefficients S_{ik} and S_{ki} were plotted versus frequency and analyzed to evaluate and confirm reciprocity and repeatability of measured data.

An important step in construction of the S_m matrix is a numeration of the antenna elements in the array. This determines the extent of symmetry of the S_m matrix and makes it possible to control the accuracy of matrix calculation. Basically because of geometrical symmetry and reciprocity the S_m matrix should have a great deal of symmetry. Symmetry problem is created by orthogonal dipoles. For the numeration purposes array is divided into two orthogonal sub arrays consisting of North-South and East-West elements. After numeration of the NS array the EW array is rotated 90° counterclockwise and is continued to numerate in the same order as the NS. The numeration scheme is shown in the Figure 1. It was found that constructed in this manner, S_m matrix has equal first and third (25x25) quadrants and the second (25x25) quadrant represents orthogonal dipole's mutual couplings. It should be noted that this feature is valid under the assumption that dipoles having the same central symmetry have equal coupling. The coupling coefficients S_{ik} were analyzed to confirm this. It was found that differences exist but are very small and averaged S_{ik} may be used to reduce the computational error. The ANA measurement errors and antenna construction tolerances are the likely source of the noted difference.

To automate S matrix creation from measurement files, a special spreadsheet processor procedure was written. For each specified frequency the procedure reads measured partial S_{ik} coefficients from files and generating of S_m matrix. The completed S_m matrix is sent to universal mathematical package for data processing.

The second step is de-embedding of the measurement hybrids from S_m matrix in order to get array S_a matrix. In present frequency band, the

antenna impedance varies over a very wide range and most accurate de-embedding of N 2-port blocks may be performed by simultaneous de-embedding operation of all 50 blocks. The de-embedding expression were defined as the following:

$$S_a = (s_{h22}S_m - \det S_h \cdot I)^{-1}(S_m - s_{h11}I);$$

Here S_h is the 2-port hybrid S matrix, I is 50x50 identity matrix. All hybrid blocks are considered having the identical S matrix. The result of the de-embedding is 50x50 matrix file S_a . The number of files generated is equal to the number of measurement frequencies.

Evaluation of the Array Performance

To accomplish feed system optimization the following performance parameters are of interest:

- active dipole impedances,
- transmitter load impedances,
- total feed system mismatch losses

To obtain the active antenna impedances, a computer procedure simulating equal incident wave array excitation was prepared. The first iteration AMU network was synthesized using the computed active impedances. It should be noted that active impedances obtained are normalized to 50 Ohms, which in general is not the AMU output impedance. To eliminate this error iterative AMU optimization by consequent AMU active impedance calculation and AMU synthesis was used.

The embedding expression for a 2-port AMU is defined as the following:

$$S_m = (s_{AMU11}I - \det S_{AMU}S_a)(I - s_{AMU22}S_a)^{-1}$$

where S_{AMU} is the S matrix of the AMUs.

The averaged antenna impedance and corresponding AMU impedance seen by transmitter are shown on Figure 3. The results show that the DP array may be effectively operated up to 8.4 MHz provided that there is no significant main lobe deterioration at high frequencies.

In a wide-band array as small as the DP, it is difficult to accomplish good matching for all antenna elements, especially under scan conditions. To estimate the impact of this effect under

operating conditions, a computer simulation for the 5x5 array was performed and the array efficiency was calculated. Efficiency is defined as ratio of total power delivered to the total incident power applied the AMU inputs. The computational expression is:

$$\eta = \frac{a^+ \cdot [1 - S_{AMU}^+ \cdot S_{AMU}] \cdot a}{a^+ \cdot a},$$

where a is a vector of incident wave.

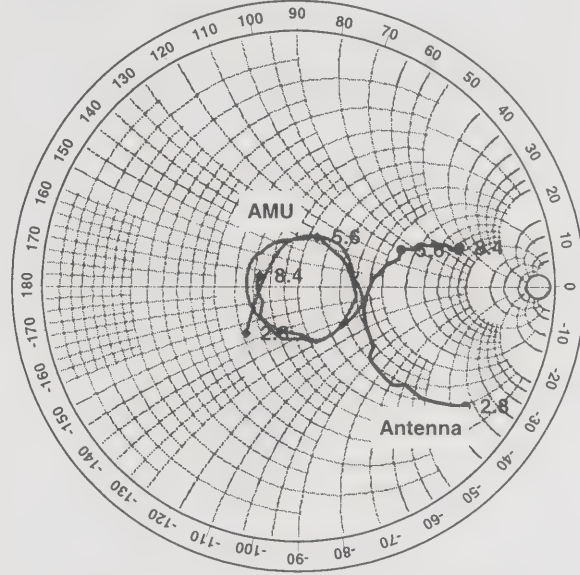


Figure 3. Average Antenna and AMU impedances

The efficiency calculations are shown in Figure 4. It should be noted that as the DP array expands beyond 5x5 size, better performance is expected.

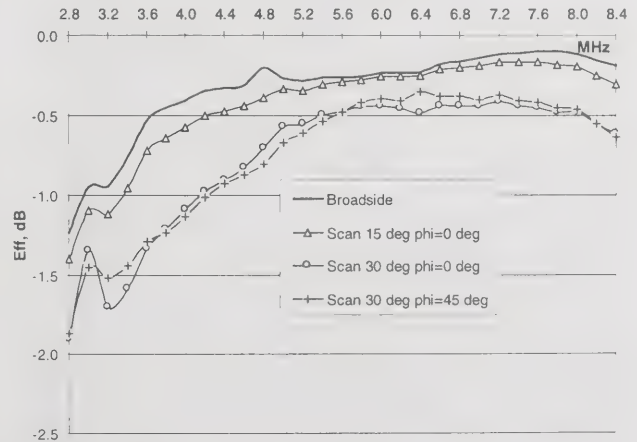
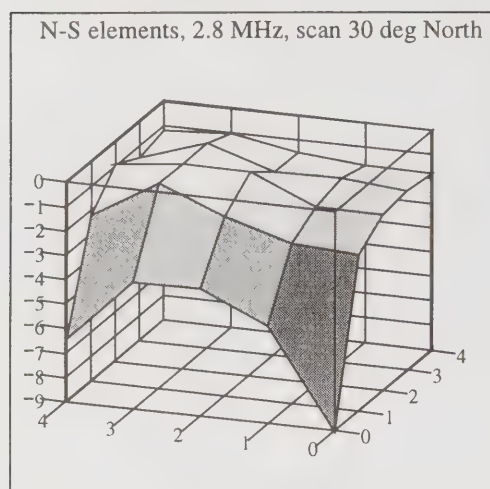


Figure 4. Array Efficiency

Considering efficiency it is also interesting to note that

- An attempt to reduce mismatch losses by designing a separate AMU network at each element would only yield an improvement in overall efficiency of 1.0-1.5 dB depending on scan angle.
- Only corner and edge elements are greatly mismatched during scan as, illustrated by Figure 5. Most other elements, including cross elements, do not experience such severe load variations.



η_{NS_dB}

Figure 5. Efficiency for N-S Elements, dB

Reference

- 1) Wideband Phased Array For Ionospheric Research Applications, 1994, Allerton Conference, Champaign, IL

Conclusions

1. A practical approach to array characterization using separate measurements of odd and even mode S-matrix is demonstrated for a wideband 5x5 cross dipole array.
2. A new AMU design with improved performance was verified.
3. The upper limit of the operating frequency band may be extended 30% from 7.0 up to 8.4 MHz.
4. Satisfactory array mismatch losses are predicted. The overall performance is within 1.5 dB of its theoretical maximum.

Advanced Chip-in-Board Thinfil Hybrid Technology with Integrated Polyimide Film Capacitors for Ultra High-Speed and Millimeterwave Applications

Wolf-Dieter Nohr*, Gerhard Hanke, Dieter-Jürgen Weber

Deutsche Telekom, Technologiezentrum Darmstadt

Agastaße 24

12489 Berlin, Germany

*+49-30-67083591, +49-30-67082161(fax), nohr@07.bln0.telekom400.dbp.de

Abstract

The processing of extreme high bit rates ($\geq 20\text{Gbit/s}$) or signal frequencies ($\geq 20\text{GHz}$) is of increasing importance for advanced telecommunications. Therefore a suitable circuit arrangement is necessary. Problems like frequency dispersion of microstrip lines, frequency dependence of parasitic impedances, and transforming characteristics of any finite length of layer must be solved. For this purpose a chip-in-board embedding technique with integrated film capacitors was developed. The contribution describes the manufacturing of such thinfil hybrid circuits containing planar capacitors and feed-through capacitors. The chip interconnection technology is represented, too.

Key words: Chip-in-board (CIB) technology, thinfil hybrid technology, reverse beam-lead technique, integrated planar capacitor, integrated feed-through capacitor

Introduction

The paper describes a new technology of assembling ultra high-speed or millimeterwave thinfil hybrid circuits. For this a new technique of inserting and innerlead bonding of chips in film circuits was developed. This technique, which we call "Reverse Beam-Lead Technology" (RBL), is suited, apart from its generally favourable applicability to the setup of MCM (Multichip Modules), especially for the implementation of ultra high-speed and millimeterwave circuits. The basic principle of the chip interconnection technology presented is that the chip to be inserted is mounted flush with the film circuit surface into a previously prepared chip-sized substrate cutout, into which interconnection-leads as a part of the film circuit project. The chips are interconnected face down from the rear side of the hybrid circuit with a flip-chip bonding device. This mounting and interconnection technology is a guarantee of nearly lossless and non-reflecting interconnections between dice and microstrip circuit.

Additional the manufacturing of integrated planar capacitors and integrated feed-through capacitors needed for coupling or bypassing is shown. For this purpose a photosensible polyimide is used as dielectric applied by spin coating (for planar capacitor) or suck applying (for the feed-through capacitor). It is possible to achieve an geometric exact adaption of the integrated capacitors to the pattern of the microstrip circuit.

To demonstrate these technologies some examples of hybrid circuits are given.

Chip Mounting

To manufacture a MIC device (= Microwave Integrated Circuit) suitable for ultra high-speed applications we use a polished thinfil Al_2O_3 -ceramic substrate (e.g. 1" x 1" x 0,01"). At first chip-sized substrate openings are cut out by using a Nd : YAG laser device. Subsequently the substrate is coated by sputter deposition with: Cr[10nm]/Au[0,2 μm].

In order to manufacture leads, which shall project into the substrate opening, the cutout has to be filled flush with suitable material. Indium proved to be suitable for that purpose. Due to its good ductility, indium can simply be brought into the substrate openings.

As a next step the desired hybrid circuit structure is created by selective galvanic plating of gold (thickness: 10 μm). For this a positiv resist mask is generated before. The leads for interconnecting the MMIC's (= Microwave Monolithic Integrated Circuits) are overlapping the filled substrate cutouts. Thereafter the indium-filling is removed with a corrosive.

GaAs-MMIC's are equipped with gold pads and so they are appropriate for mounting by soldering. For this purpose a solder is deposited on the underside of the leads. We use indium by galvanic plating from a commercial sulphamate electrolyte. The thickness of

the solder layer is about $10\text{ }\mu\text{m}$. A cathodic current density of 1.5 A/dm^2 is proved to be advantageous. After remelting of the indium solder layer the leads are provided with "soft bumps". With that a MIC prepared for chip embedding is created. In figure 1 from the rear side of the MIC a substrate cutout with projecting leads provided with indium bumps are shown.



Fig.1: Substrate cutout (1mm x 1mm) with indium-bumped leads (width: $100\mu\text{m}$).

The MIC is placed face down on a programmable hot-plate to carry out the soldering process. The chip is inserted face-down into the chip-sized cutout and interconnected by soldering. Figure 2 shows the principle of face-down mounting of a chip using a component placer.

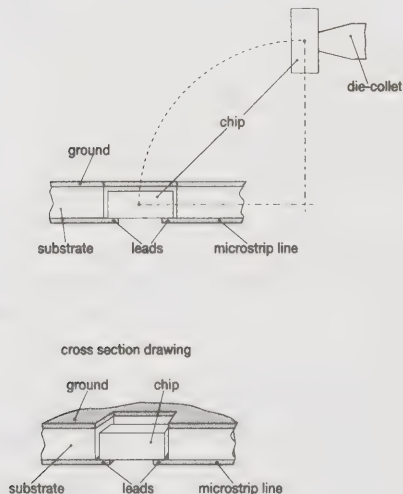


Fig. 2: Principle of CIB face-down mounting.

To interconnect Si-chips - equipped with aluminium pads - with Au-leads we investigate the method of simultaneous ultrasonic welding. For this purpose the face-down fineplacer was equipped with an ultrasonic-generator.

This technology can be advantageous for interconnecting Au-leads with Au-pads (e.g. GaAs-chips), too. To receive welding interconnections with high reliability the leads are provided with gold bumps ("hard bumps"), shown in figure 3.

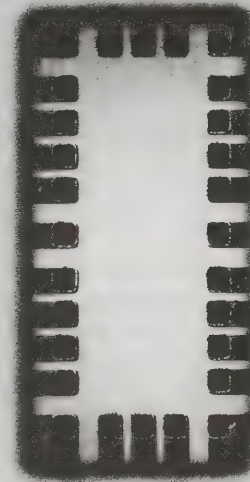


Fig. 3: Substrate cutout (1mm x 2mm) with gold bumped leads (Width: $100\mu\text{m}$).

Capacitors

To manufacture integrated capacitor devices also the thin film technology is used. After producing of the basic electrode by photolithographic structuring/additive galvanic plating (Cr/Au//Au) the dielectric layer is realized. For this purpose a photo-sensible polyimid is used as dielectric applied by spin coating (for planar capacitor) or suck applying (for the feed-through capacitor). The dielectric constant is about $\epsilon_r = 3.3$. As next step the top electrodes are produced. For it the semi-additive technology is used, too (Cr/Au//Au). After forming the top electrodes the hybrid circuit is complete. In figure 4 the photograph of a planar capacitor is shown. Onto basic electrodes as part of the microstrip line the dielectric is generated. Over it the top electrode is formed. Figure 5 shows a photograph and a measured S_{21} -parameter diagram of a feed-through capacitor up to 50 Ghz. In this case the capacitor is assembled inside a through hole. The basic construction is identical with a planar capacitor. In order to achieve a capacitance of such blocking capacitor of $C = 2...3\text{pF}$ it is necessary to apply a thin very even polyimid film. The thickness has to be about $3..5\mu\text{m}$.

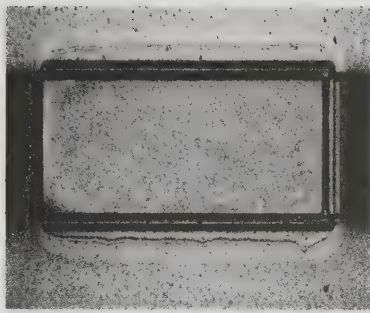


Fig. 4: Planar capacitor as part of microstrip line (width of the strip line: 254 μ m).

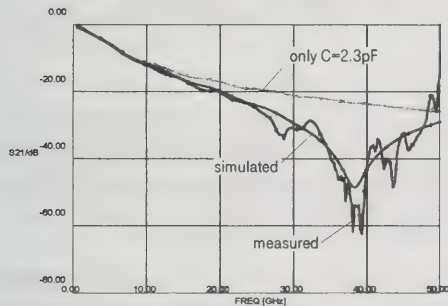


Fig. 5: Integrated feed through thin film capacitor (diameter: 300 μ m) and measured S₂₁-parameter diagram up to 50 GHz.

The high frequency characteristics of the coaxial capacitors were measured with the aid of an S-parameter equipment (HP 8510/8517). They were placed in the middle of an 1 inch long 50 Ω microstrip line. The calculated equivalent circuit can be described roughly by the series connection of $C = 2.3$ pF, $R = 0.1$ Ω and $L = 7$ pH. Because the inductance is given by length and diameter of the drilling the resonance frequency must be influenced by the value of the capacitance, what can be done rather easily by varying the thickness of the dielectric material. The measured characteristic of a real coaxial capacitor is compared with the characteristic of the equivalent circuit and an ideal capacitor of 2.3 pF. The resonance near the bit repetition rate gives an additional and wanted effect for the shunting quality of this component.

Circuits

Figure 6 shows a RBL-mounted 20 Gbit/s laser driver. The laser driver is fabricated in GaAs-HEMT-technology by Fraunhofer-Institut für angewandte Festkörperphysik, Freiburg, Germany. You can recognize the exact adaption of microstrip lines of the MIC to the chip-pads.

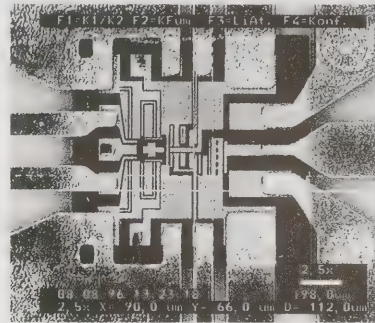


Fig. 6: RBL-mounted 20 Gbit/s laser driver chip (chip-size: 900 μ m x 900 μ m).

In figure 7 a 40 GHz small band power amplifier (HP HMMC-5040) for use as clock driver in a 40 Gbit/s OTDM-system is shown. Hybrid integrated thinfilm bypass capacitors are used, the capacitance is about 5pf/mm². Due to fabrication tolerances of the GaAs-chips the air gap between chip and ceramic substrate has to be 50 μ m.

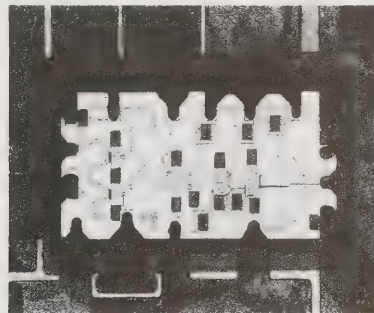


Fig.7: Reverse Beam-Lead mounted 40 GHz amplifier with planar capacitors (chip-size: 750 μ m x 1650 μ m).

Figure 8 shows a RBL-mounted 40 Gbit/s 2:1-multiplexer. The GaAs-chips were developed by Fraunhofer-Institut, Freiburg/Germany and use a 0.2 μ m source coupled FET logic. Since all broadband inputs and outputs are of differential type, coupled microstrip lines can be used to reduce the required space for the fourteen 50 Ω lines around the chip,

see details in Fig. 8. Bypassing of power supply connections is done by hybrid integrated surface capacitors and by coaxial feed-through capacitors.

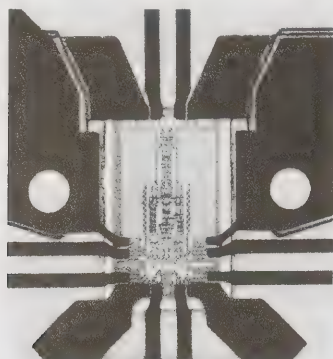
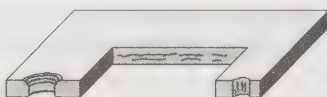


Fig.8: RBL-mounted 40 Gbit/s 2:1-multiplexer with planar and feed-through capacitors (chip-size: 980 μ m x 1750 μ m).

Figure 9 gives a technological overview of the reverse beam-lead technique with integrated capacitors:

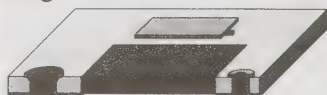
1. Polished Al₂O₃-substrate with cutouts:



2. Ground and basic electrodes:



3. Indium filling:



4. Dielectric:



5. Indium filling; top electrodes, microstrip lines:



6. Indium removal:



7. Chip mounting:



Conclusions

The presented simultaneous CIB interconnecting technology is an effective tool for ultra high-speed and millimeter wave applications. We obtain nearly lossless and non-reflecting interconnections between chips and microstrip lines and also very low impedance ground and power supply interconnections. Comparative measurements with conventional bonded circuits up to 50 GHz show the advantage of the new technology.

The optimum results will, however, be obtained by a more sophisticated metallization layout of high-frequency and high-bitrate chips. The impedance of connecting lines on the substrate should continue on the chip. The chips should be cut very exactly to get a gap free fit between chip and substrate cutoff. We are sure that our technique will be applicable also at frequencies above 50 GHz. This technology still includes more advantages than we could describe in this short paper. We are working in many directions and for many different applications. Further investigations will be concentrated, particularly, on improving the reliability of the described simultaneous interconnecting technique.

References

- [1] W.D. Nohr and G. Hanke: "Reverse Beam-Lead Interconnections for Ultra High-Speed Multichip Applications", Proc. 5th Intern. Conf. and Exhibition on Multi-chip Modules, Denver, CO, April 1996.
- [2] G. Hanke: "20 Gbit/s Experimental Transmitter and Receiver for Fiber Communication Systems", Proc. IEEE Global Telecommunications Conference, Houston, TX, November 1993, pp. 1144-1148.
- [3] W.D.Nohr, G.Hanke: „Chip-in-Board (CIB) Aufbautechnik für Ultra High-Speed und Millimeterwellenanwendungen"; Proc. 11. Internationale Messe und Kongreß für Systemintegration; Nürnberg/Germany; Apr. 1997.
- [4] Berroth, e.a., „20-40 Gbit/s 0.2 μ m GaAs HEMT Chip Set for Optical Data Receiver"; Proc. of GaAs IC Symposium 1996, pp 133-136.
- [5] Lang, M. e.a., „GaAs HEMT IC for 40 Gbit/s Data Transmission Systems"; Proc. of EMAC, Barcelona, May 1997.
- [6] Nohr, W.-D., Lemanski, H., Meusel, E., „Simultaneous CIB Interconnection Technology for Ultra High-Speed Applications."; Proc. of 11th European Microelectronics Conference, Venice / Italy, May 1997, pp. 558-561

Characterization of embedded resistors for high frequency wireless applications

Nanju Na, Kwang Lim Choi and Madhavan Swaminathan
Georgia Institute of Technology

777 Atlantic Drive, Atlanta, GA30332

404-894-3340, 404-894-9959(Fax), madhavan.swaminathan@ee.gatech.edu Abstract

Abstract

This paper analyzes frequency and time domain response of embedded resistors in a multilayered structure using LTCC (Low Temperature Co-fired Ceramic) technology and discusses their effect on high frequency operation of mixed signal circuits using them. The results have been compared to TDR (Time Domain Reflectometry) and TDT (Time Domain Transmission) measurements using pulse propagation. Frequency-dependent parasitics are observed to be significant in embedded resistors from measurements and simulations. The importance of accurate modeling including frequency dependence and parasitics is emphasized in this paper for prediction of the frequency behavior of embedded resistors. Macromodels for resistor structures were developed and their simulation results show a good agreement with measurements. This paper also discusses a macromodeling approach that provides good correlation with measurements.

Key words: embedded resistor, macromodeling

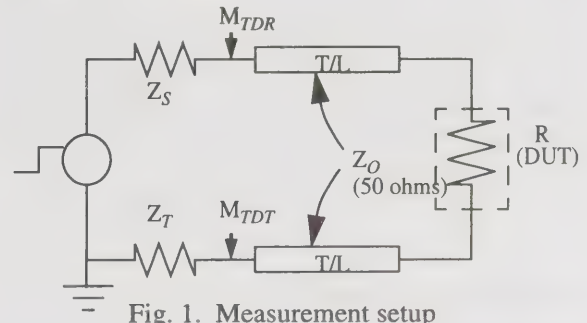
1. Introduction

The trend in portable wireless electronics is to combine digital and RF circuits into a compact packaged mixed signal module. Examples of such electronics are pagers, cellular phones, transceivers, and global positioning systems that function in the frequency range DC-5GHz. Embedded components which is an emerging technology area provides a method for achieving size shrinkage by replacing surface mount components in a multi-layered packaging environment. The embedded components, however, have different electrical characteristic as compared to discrete components due to their geometric structure and parasitic effects. Therefore, these effects should be characterized and carefully considered in mixed signal wireless applications. In this paper embedded resistors have been characterized using a combination of measurements and simulation. A novel technique based on macromodeling has been used and compared to TDR and TDT measurements. A resistor test vehicle with 20 resistors embedded in a 4 layered substrate using LTCC-M technology was used for the study.

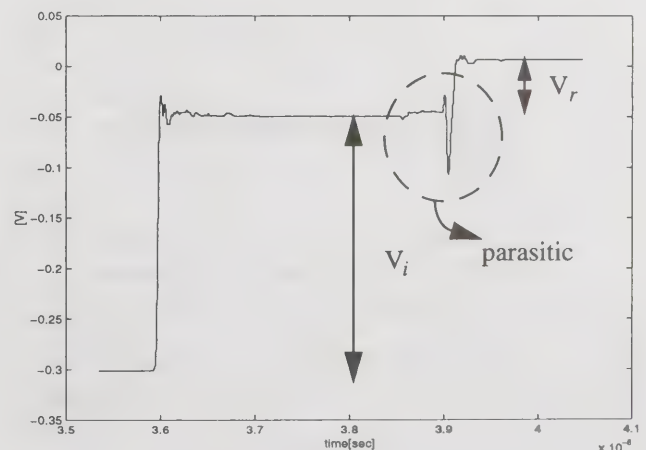
2. Measurement and DC Resistance Calculation

The measurement was set up as shown in Figure 1 where R is the resistor under test and Z_S and Z_T are terminating impedances of the digital sampling oscilloscope. A pulse with transition rise time of 35ps and bandwidth of 20GHz was used as the incident wave in Figure 1. The resistance values calculated from the measurement showed good agreement with the specification values within 5% error. Parasitic capacitances and

inductances were observed from the TDR measurements for most small resistors.



Similarly, frequency dependence was seen for most large resistors. Parasitics and frequency dependence in an embedded resistor are a function of the size and physical geometry of their structure. For embedded



resistors in a multilayered structure, the effect of capacitive parasitic can become significant due to a small distance between the resistor and ground plane. In the measurement of the resistors, parasitic capacitance was observed to be the most significant for resistors with large widths (small resistance values) and the frequency dependence was prominent in resistors with large length (large resistance values). Figure 2 shows an example of TDR waveform measured for the resistor with value 26.04Ω at M_{TDR} . Peaks and valleys at TDR measurements imply the parasitic inductance, capacitance or impedance mismatch.^[1] From Fig. 2, the parasitics associated with the device can be clearly seen, showing the dominating capacitive parasitic.

3. Developing the Macromodels

The resistor macromodels were developed for 26.04Ω and 858.43Ω which were considered as reasonable samples having the lowest and the highest resistances respectively in the test vehicle. Though not compared with measurements, a macromodel for 50Ω resistor was developed for transmission line analysis. It is important to note that simple resistor models give inaccurate results due to the existence of parasitics and frequency dependence in the structure. Figure 3 shows a resistor structure analyzed in SONNET in order to extract the frequency response data (S parameter data).

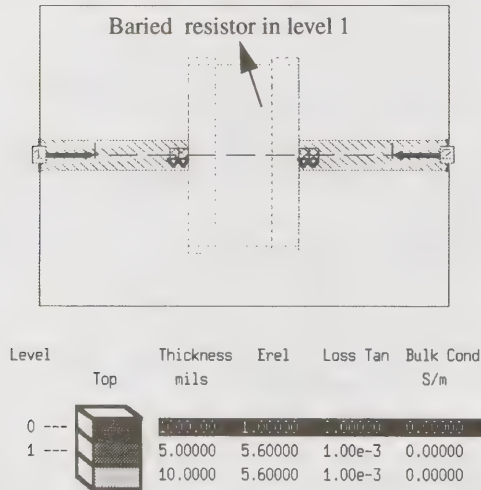


Fig.3 The physical structure analyzed in SONNET: Top view(upper) and the layer information

A novel technique^[2] based on Cauchy's method using the data obtained from SONNET^[3] was used to extract a polynomial representation for the 2-port y parameters in the form

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (1)$$

where

$$y_{ij}|_{i=j} = \frac{a_0 + a_1s + a_2s^2 + a_3s^3}{d_0 + d_1s + d_2s^2 + d_3s^3 + d_4s^4}$$

and

$$y_{ij}|_{i \neq j} = \frac{b_0 + b_1s + b_2s^2 + b_3s^3}{d_0 + d_1s + d_2s^2 + d_3s^3 + d_4s^4}$$

In (1) all the parameters have the same poles and provide excellent results from DC to 4GHz. To create SPICE macro models, g (voltage controlled current source) elements were used. Figure 4 shows y_{11} and y_{21} of 858.43Ω resistor generated by SONNET, and the rational polynomial representation. In eq.(1), 4 left half plane poles have been used to represent the y parameters. From the order of this model, time/frequency response of the resistor is expected to be that of an equivalent circuit containing at least 4 reactive passive components which are frequency-dependent.

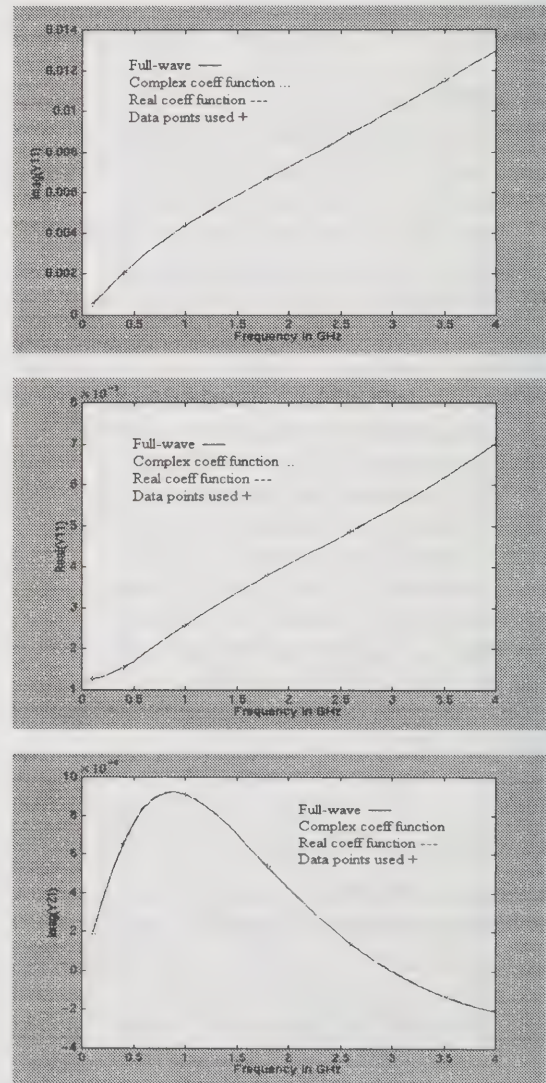


Fig. 4 Transfer function for 858.43Ω resistor: $\text{Im}(y_{11})$, $\text{Re}(y_{11})$, and $\text{Im}(y_{21})$ from the top

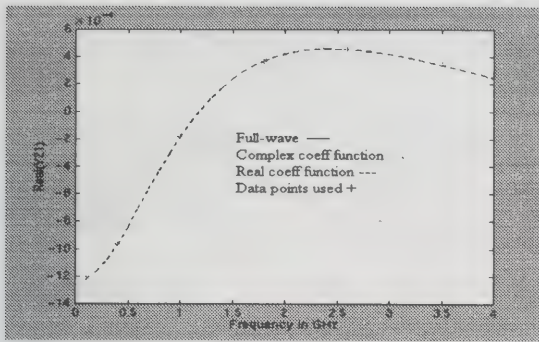


Fig. 4 (Continued) $\text{Re}(y_{21})$ 858.43 Ω resistor

4. Measurement and Spice Simulation Comparison

TDR and TDT measurements were simulated by modeling every node of the measurement setup including time delays and impedances in cables and probes in spice. Two approaches were used to develop spice models for the resistors and compare the results. The first approach used a simple resistor model for the embedded resistor. This approach is often used but can produce erroneous results as shown in the simulation result. The second approach used the macro models developed above with all the parasitics and frequency-dependence included.^[2,4] Figure 5 and Figure 6 show the TDR measurements and their simulation results for 858.43 Ω resistor and 26.04 Ω resistor respectively.

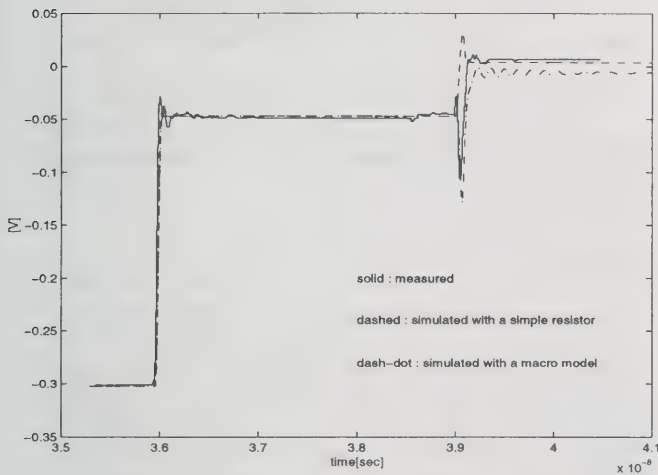


Fig.5 TDR measurement and simulations for 26.04 Ω

Small positive peaks shown in the simulation using the simple resistor model is due to the high impedance ($>50 \Omega$) transition in the test vehicle. The simulations using macromodels show good agreement with measurements in all cases as compared to the simple resistor model, clearly revealing frequency dependence and parasitics. The effect of rounding of the pulse due to high

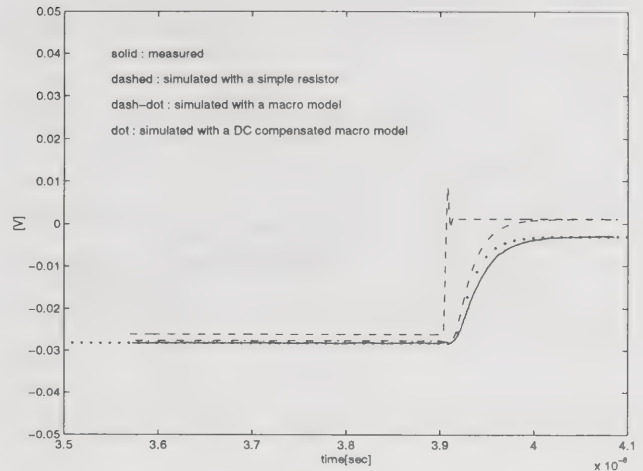


Fig.6 TDT measurement and simulations for 858.43 Ω

frequency effects can be clearly seen in Figure 5 for the 858.43 Ω resistor which has been captured accurately by the macromodel. The rounding and slanting of the pulse introduces appreciable delay at the 50% level. The 5 mV difference in DC levels between measured values and simulated values in TDT measurement can be attributed to the DC resistance. The modeled resistance value of 858.43 Ω (spec.) was measured to have 863 Ω using LCR meter. Using DC compensation by adjusting the a_0 , b_0 , and c_0 terms in eq.(1), the agreement between model and measurement was improved as shown in Figure 6.

5. Impact of Embedded Resistors on Circuit Behavior Using Simple Networks

To study the effect of terminating transmission lines using embedded resistors, a 50 Ω transmission line terminated with a 50 Ω embedded resistor was simulated using a step with a 35 ps rise time. Due to the capacitive and inductive nature of the resistor, the ringing in the reflected response can be seen in Figure 7 which can cause problems in high speed circuits. Next, voltage divider networks using two embedded resistors were simulated using a pulse, as shown in Figure 8 to 10. In Figure 8, the 858.43 Ω resistor divider network causes a considerable slow-down of the rise time while the 26 Ω resistor network causes oscillation based on the rise time (35ps or 250ps). Thus considerable emphasis in designing embedded resistors is necessary for high speed circuits.

6. Conclusion

The embedded resistors were measured and the associated parasitics were compared from the TDR measure-

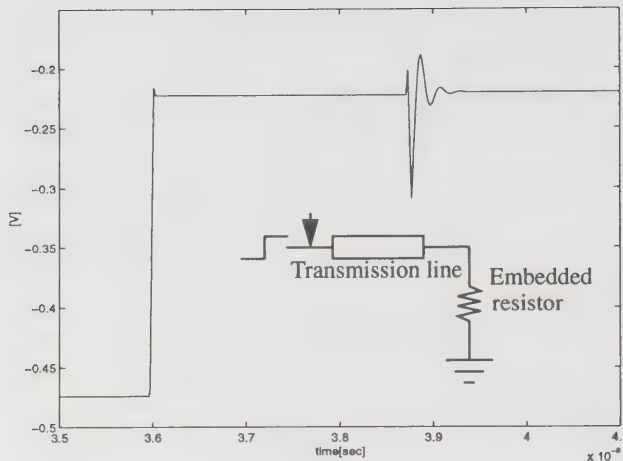


Fig.7 Transient response of a transmission line terminated with matching impedance of embedded resistor

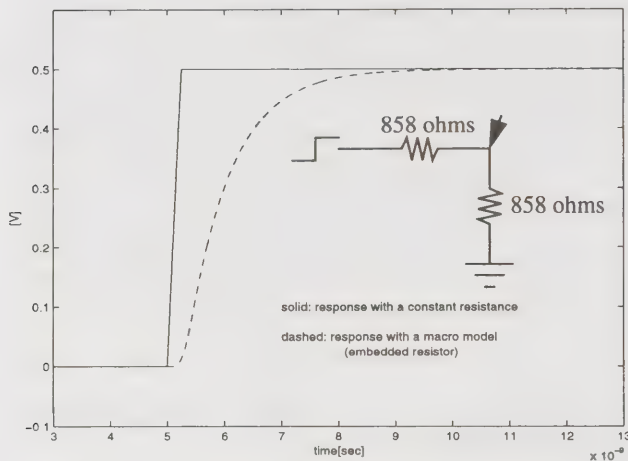


Fig.8 Transient response of 858 ohms resistor

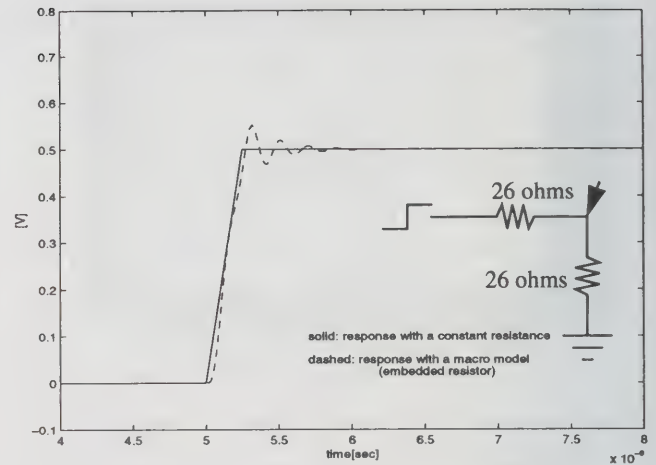


Fig. 9 Transient response of 26 ohms resistor for a step input with 250 ps rise time

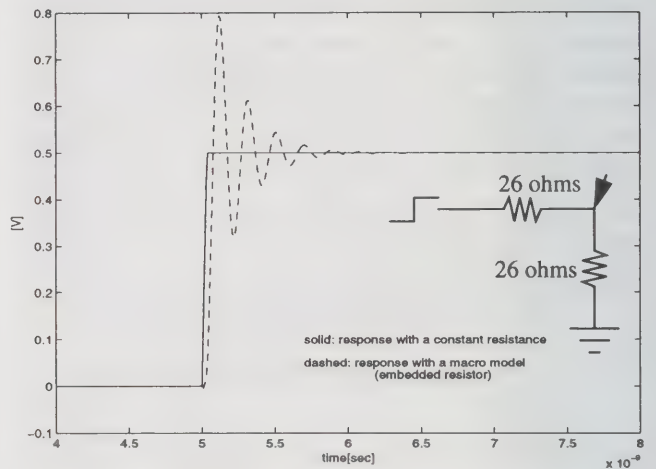


Fig.10 Transient response of 26 ohms resistor for a step input with 35 ps rise time

ment. General observation showed that the parasitic capacitance for small value resistors (large width) is the most significant parasitic in embedded resistors and the inclusion of frequency dependence is critical for large value (long) resistors.

A novel method based on macromodeling was used to simulate the embedded resistors in spice and the results were compared with simple resistor models. The simulation with a macro model showed good agreement with measurements indicating the importance of parasitics and frequency dependent responses. The importance of parasitics and frequency dependence have been discussed for simple termination and voltage divider networks.

References

- [1] Richard E. Matick, *Transmission Lines for Digital and Communication Networks*, McGraw-Hill.
- [2] Kwang Lim Choi and Madhavan Swaminathan, "Utilization of Fast Algorithm to Analyze Embedded Passive Components Using Commercial EM Solvers", IEEE 6th topical meeting on Electrical Performance of Electronic Packaging, pp.240-243, Oct. '97.
- [3] Sonnet User's Manual Vol. 1 & 2, Release 4.0, 1996.
- [4] Anisha Sood, Kwang Lim Choi, Anand Haridass, Nanju Na, and Madhavan Swaminathan, "Modeling and Mixed Signal Simulation of Embedded Passive Components in High Performance Packages", pp. 506-511 MCM Conference (Multi-Chip Module), April '98.

A PHEMT MMIC Broad-Band Power Amplifier for LMDS

Young-Gi Kim, Sung-Jae Maeng*, Jin-Hee Lee*, and Chul-Soon Park*

Dept. of Data Communication, Anyang University, 707-113, Anyang 5-Dong, Manan-Gu,

Anyang-City, Kyungki-Do, 430-714, KOREA

+82-343-67-0894(fax), kimyg@aycc.anyang.ac.kr

* Dept. of Compound Semiconductor, Electronics and Telecommunications Research Institute

161 Kajong-Dong, Yusong-Gu, Taejon, 305-350, KOREA

+82-42-820-6183(fax), sjmaeng@etri.re.kr

Abstract

A two-stage monolithic microwave integrated circuits (MMIC) broad-band power amplifier with AlGaAs/InGaAs/GaAs pseudomorphic high electron mobility transistor (PHEMT) has been developed for the up-link and down-link applications for local multipoint distribution systems (LMDS) in the frequency range of 24 ~28 GHz. The amplifier has a small signal gain of 18.6 dB at 24.5 GHz and 16.7dB at 27.1GHz. It achieved output powers of 19.8 dBm with PAE of 19.8% at 24.5 GHz and 18.8 dBm at 27.1 GHz.

Key words: power amplifier, MMIC, broad-band amplifier, PHEMT, LMDS, Ka-band, driver amplifier

Introduction

As demand for high local access speed and large capacity grows, Ka-band local multipoint distribution systems (LMDS) become very attractive for future communications. The service radius of the LMDS cell, which is very closely related with cost, is mainly limited by the power output of Ka-band technology [1].

Several leading groups have demonstrated their recent developments of the Ka-band power amplifiers based on pseudomorphic high electron mobility transistor (PHEMT), heterojunction bipolar transistor (HBT) and metal semiconductor field effect transistor (MESFET) technology [2]-[4]. Because of its excellent transport properties, PHEMT is ideally suited for MMIC amplifier implementation for high-efficiency, high-gain application in Ka-band frequencies. Ion-implanted MESFETs have been proven to produce cost savings from material and processing stand point [5]. Monolithic microwave integrated circuits (MMIC) have been widely investigated because they have advantage in terms of miniaturization, good reproducibility, and low cost in high production, without undesirable wire connection that strongly impacts assembly yield and radio frequency performance up to millimeter-wave range, which is very profitable for LMDS circuit [6].

The LMDS needs up-link and down-link

wireless connections with separate frequency bands if it is to be used two-way communications system. It needs to be kept some distance between these two bands in order to avoid frequency interference. The up-link radio frequency (RF) power is much smaller than that of the down-link. Most of the research works have been focused on maximum output power handling capacity for the Ka-band LMDS down-link communications. Design and production of the down-link driver amplifier and up-link power amplifier are unavoidable. These additional works are not efficient in economic point of view due to its relatively low volume demand in the early stage of LMDS evolution. This paper presents a multi-purpose broad-band MMIC power amplifier that can be used for the up-link power amplifier and down-link power amplifier driver circuit.

Design

For the broad band amplifier circuit, distributed or negative feedback scheme can be used. The former can produce very broad band, up to 23 GHz band width, but the degradation of the power gain and output power is very large. Therefore, the former needs larger active devices and wider area of the wafer, due to larger number of the devices, than the latter when the MMIC amplifier is formed for to produce some amount of output power and gain [7].

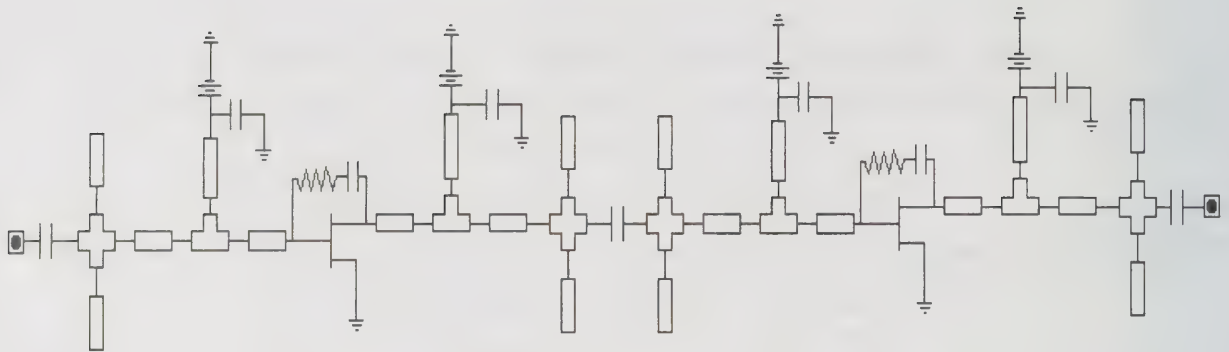


Figure 1 Schematic diagram of a two-stage broad band power amplifier

The introduction of negative feedback by adding a resistor network from the output to the input helps improving the transistor stability. In addition, the effect of feedback is to make the input and output impedance more convenient for matching [8].

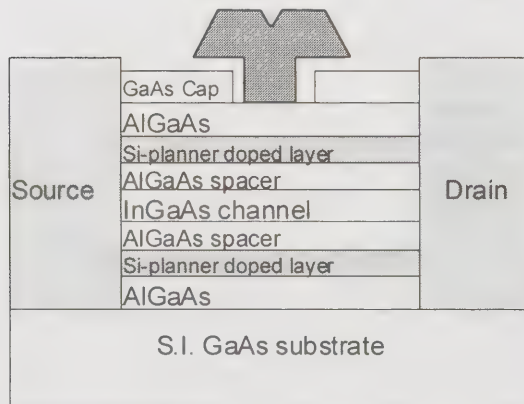


Figure 2. Cross-sectional view of the PHEMT

A two-stage amplifier was designed as shown in Figure 1. Resistive feedback from the drain to the gate of the PHEMT was used for broad band match and stable amplification. AlGaAs/InGaAs/GaAs PHEMT device structure grown by MBE with T-shaped gate of 0.15 μm gate length, shown Figure 2, was used for this work. The PHEMTs with gate width of 200 μm and 400 μm , were used to produce

10 dB amplification in the first stage amplifier and 8.5 dB amplification in the second stage respectively. Both amplifiers were designed to class A operation using micro-strip lines. Symmetric open stub conjugate matching scheme was applied to prevent an unexpected oscillation. MIM capacitors were used only for coupling and decoupling purposes to reduce process variation. The bias network consists of a high impedance quarter-wave transmission line with decoupling capacitor serving as RF short circuit.

Output power, flat power gain over band width, matching, stability, efficiency, and chip size were optimized for the up- and down-link power applications of the LMDS. LIBRA microwave circuit analysis program with Root FET model was used for the simulation.

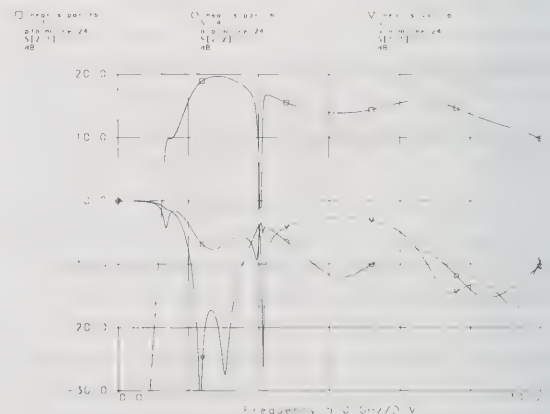


Figure 3 Simulated Small signal performance of a broad band amplifier.

Simulated linear gain, input broad band matching and output broad band matching of designed circuit are shown in Figure 3.

Output powers and power added efficiencies (PAE) versus input powers were simulated for 24.5 GHz and 27.1 GHz as shown in Figure 4.

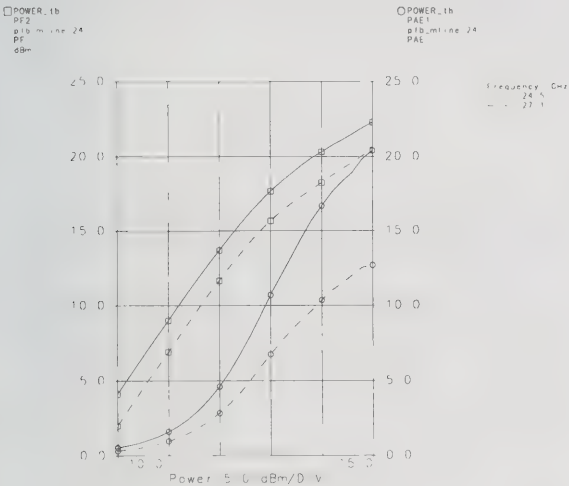


Figure 4. Simulated output powers and PAEs versus input powers for 24.5 GHz and 27.1 GHz

Fabrication

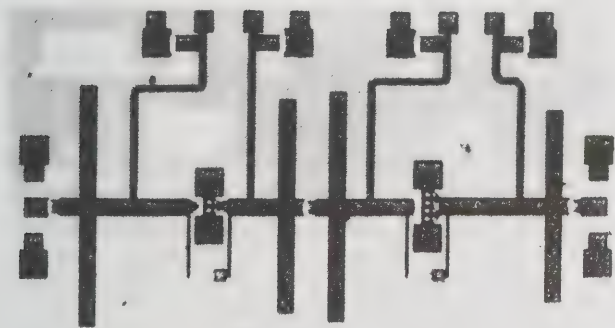


Figure 5. Photograph of two-stage broad band monolithic amplifier

The power amplifiers were fabricated by standard PHEMT process. The wafer was thinned to 100 μm . Back-side via holes were etched to the ground pad of the device to provide a low source inductance and high thermal dissipation. Then the

wafer was gold-plated. Figure 5 shows a fabricated chip. The chip size is 1.44 mm x 2.45 mm.

Measured Performance

The two-stage amplifier had a small signal gain of 18.6 dB at 24.5 GHz and 16.7dB at 27.1GHz when biased with 3 V on the drains as shown in Figure 6. The measured gain was a little higher than that of the simulated one.

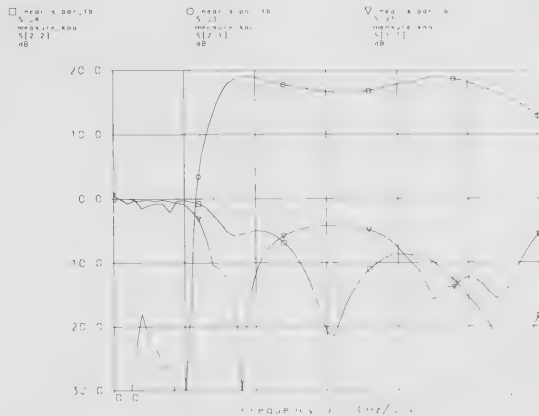


Figure 6. Small signal performance of the fabricated two-stage MMIC amplifier

Figure 7 shows output power and PAE for 0 dBm input as a function of frequency.

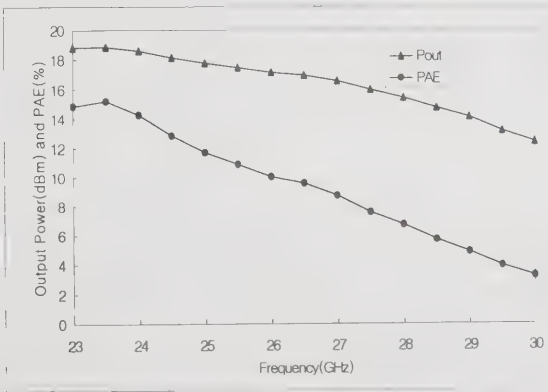


Figure 7. Output power performance and PAE for 0 dBm input power as a function of frequency

The amplifier was capable of 18.5 dBm output power with 17.5 dB gain at 24.5 GHz for up-link power amplifier as shown Figure 8. The power-added efficiency was 13.5% at this point.

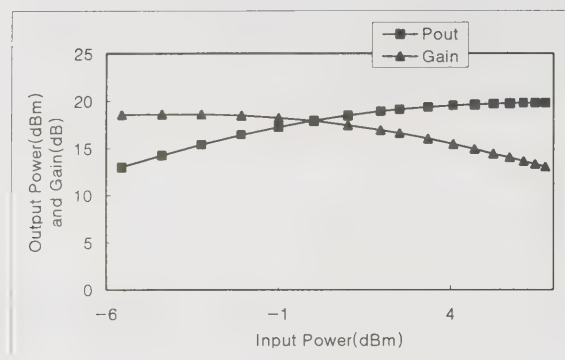


Figure 8. Output power and gain versus input power at 24.5 GHz

When the amplifier was measured for down-link power amplifier driver, it was capable of 17.5 dBm output power with 15.7 dB gain at 27.1 GHz as shown Figure 9. Corresponding PAE was 11%.

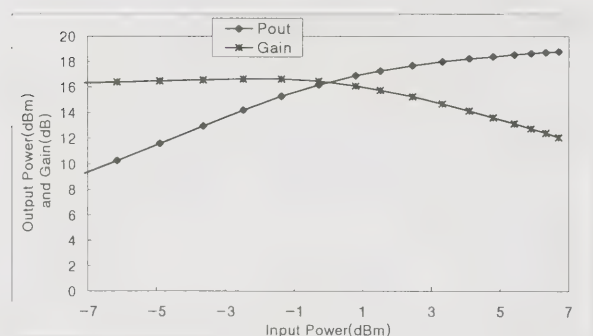


Figure 9. Output power and gain versus input power at 27.1 GHz

It achieved output powers of 19.8 dBm with PAE of 19.8% at 24.5 GHz and 18.8 dBm at 27.1 GHz when it was optimally biased.

Summary

A Ka-band multi-purpose two-stage broad-band MMIC PHEMT power amplifier with resistive parallel feedback has been demonstrated. It can be used as a down-link final stage power amplifier over 24.25 to 24.75 GHz or an up-link driver amplifier over 26.7 to 27.5 GHz for the two-way LMDS.

REFERENCES

1. Pauline Tratter and Kate Hewett, "Broadband Wireless Strategies," Ovum Ltd, 1997
2. Mansoor K. Siddiqui, Arvind. Shurma, Leonardo G. Callejo, Chung-Hau Chen, Kin Tan, and Hauan-Chun Yen, "A high power and high efficiency power amplifier for local multipoint distribution service," IEEE MTT-S Digest, pp. 701-704, 1997
3. S. Murakami, S. Tanaka, Y. Amamiya, H. Shimawaki, N. Goto, K. Honjo, Y. Ishida, Y. Ishida, M. Yajima and Y. Hisada, "A 3.6-W 26 GHz-band AlGaAs/GaAs HBT power amplifier," IEEE GaAs IC symp. Tech Dig., pp.99-102, 1996
4. Y. Kalayci, R. Tempel, W. Lutke, M. Akpinar, and I. Wolff, "A miniaturized Ka-band MMIC high-gain medium power amplifier in coplanar line technique by using a conventional 0.5 μ m MESFET technology," IEEE MTT-S symp. Dig., vol. 3, pp.1623-1626, 1995
5. Ron Yarborough, Paul Saunier, and Hua Quen Tserng, "Performance comparison of 1 Watt Ka-band MMIC amplifiers using pseudomorphic HEMTs and Ion-Implanted MESFETs," IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium, pp.21-24, 1996
6. K. W. Kobayashi, H. Wang, R. Lai, L. T. Tran, T. R. Block, P. H. Liu, J. Cowles, Y. C. Chen, T. W. Hauang, A. K. Oki, H. C. Yen and D. C. Streit. "An InP HEMT W-band amplifier with monolithically integrated HBT bias regulation," IEEE Microwave and Guided Wave Letters, Vol. 7, No. 8, pp.222-224, Aug. 1997
7. Keith M. Simon, Ratana M. Wohlert, John P. Wendler, Lisa M. Aucoin and David W. Vye, "K-Through Ka-band driver and power amplifiers," IEEE Microwave and Millimeter-Wave Monolithic Circuits Symposium, pp.29-32, 1996
8. I. D. Robertson, "MMIC design," IEE, 1995

A Two Stage, Monolithic Integrated 200 mW HEMT Amplifier for Wireless ATM

Thomas A. Bös, Urs Lott, Werner Bächtold

Swiss Federal Institute of Technology (ETH) Zürich,
Laboratory for EM Fields and Microwave Electronics
Gloriastr. 35, CH-8092 Zürich, Switzerland

Tel.: + 41-1-632 59 77, Fax + 41-1-632 11 98, boes@ifh.ee.ethz.ch

Abstract

For wireless LAN applications a compact two stage HEMT amplifier is designed and on-wafer measured with on-chip matching and bias networks. At 17.2 GHz the amplifier delivers 23.1 dBm saturated output power from a 3.3 V supply. At the 1 dB compression point of 20.4 dBm the amplifier has 13.1 dB gain. Measured performance of the amplifier proves the applicability up to 20 GHz. The amplifier having a low output reflection coefficient is suitable as driver stage for amplifiers with higher power outputs.

Key words: Wireless ATM, WLAN, power amplifier, HEMT, amplifier, Magic WAND

Introduction

High speed wireless local area networks for indoor use are becoming more and more important. WLAN systems with a data rate up to 2 Mb/s are already standardized in IEEE 802.11 [1]. New multimedia services require higher data rates. Wireless ATM is one solution for accessing wired, high data rate ATM networks [2]. To provide the high transmission capacity new frequency bands have to be allocated. The emerging European HIPERLAN standard addresses frequency spectra in the 5 and 17 GHz range [1]. In the European ACTS project "Wireless ATM Network Demonstrator" a system with data rates of 20 Mb/s in the 5 GHz range is in development and investigations of the requirements for data rates up to 155 Mb/s in the 17 GHz band have been made taking into account system architecture and technological constraints [3, 4, 5].

For portable computer application the size of the power amplifier and the supply voltage are limited. For the 17.1 - 17.3 GHz frequency range the measured results of a two stage, on chip matched amplifier with a chip size of only 1.5 mm² are shown. From a 3.3 V supply the compact amplifier delivers 100 mW output power at P_{1dB} and 200 mW at P_{sat} . Furthermore

measured performance figures up to a frequency of 20 GHz are given. This broadband performance allows the application of the designed HEMT amplifier in WLAN systems at higher frequencies, e.g. 19 GHz [5].

IC Realization

In Fig. 1 a block diagram of the two stage amplifier is shown. The transistor of the driver stage has a gate width of 300 μ m. The input matching, the gate and drain biasing network as

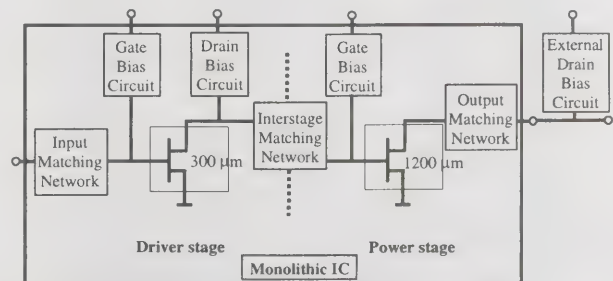


Fig. 1: Block diagram of the power amplifier

well as the interstage matching network are integrated on the chip. In the power stage the 1200 μ m transistor and its gate biasing and the output matching network are integrated. The drain bias network is external due to the high DC

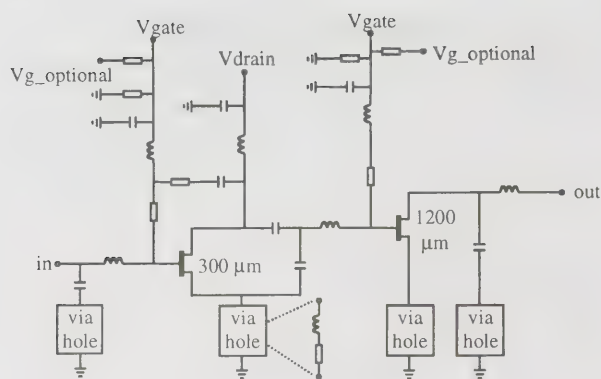


Fig. 2: Basic schematic diagram of the monolithic power amplifier

current required by the transistor. The low pass type output matching network, shown in the basic schematic diagram (Fig. 2), allows the external biasing. No harmonic tuning was done at the output. The output stage is designed conditionally stable.

The capacitor of the interstage matching network and the source of the first transistor are connected at the same via hole, which is modeled as lossy reactance. In contrast to two separate via holes,

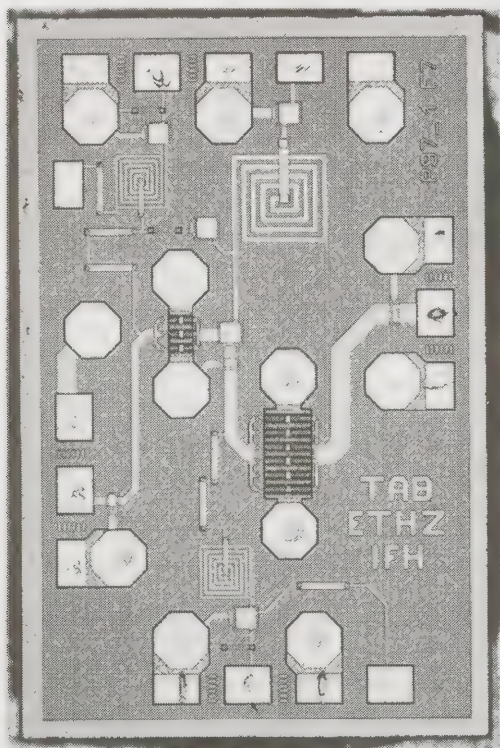


Fig. 3: Photograph of the chip (1 x 1.5 mm²)

simulated results showed that a broader interstage and input match is achieved.

Negative feedback stabilizes the driver stage. The whole circuit was simulated with a harmonic balance simulator. The HEMTs are simulated with a Statz MESFET model modified with an RC network between drain and source. Thus, inaccuracies in the large signal simulation results are expected.

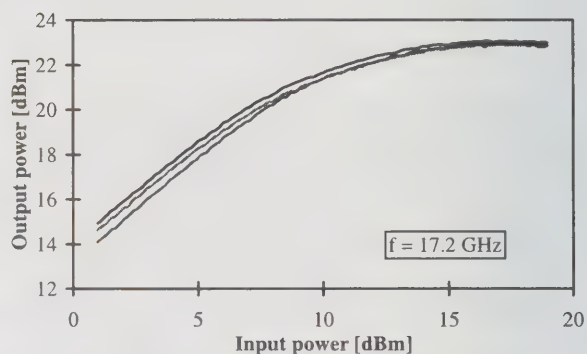


Fig. 4: Measured fundamental output power of three samples

The power amplifier was designed with a standard foundry Pseudomorphic HEMT process Philips D02AH. The transistors have a gate length of 0.2 μm with $f_t = 60$ GHz. Their fingers are sized 6 x 50 μm and 12 x 100 μm. In the RF path all inductances of the schematic are modeled and realized with microstrip line elements, whereas spiral inductors are used in the bias networks. Via holes to the back metal provide an optimum source grounding, but limit a maximum layout density by the required minimum distance between two adjacent via holes. The required center to center spacing is 200 μm. Nevertheless, the layout (Fig. 3) is very compact with a chip size of 1 x 1.5 mm².

Measurements

For all on-wafer measurements a drain supply voltage of 3.3 V was used at both stages. The gate source voltage of the first transistor was 0.0 V. The power transistor was biased at $V_{GS} = -0.3$ V. The input versus output power characteristic at 17.2 GHz is measured for three samples (Fig. 4). The saturated output power is 23.1 dBm. The output 1 dB compression point is at 20.4 dBm with a typical gain of 13.1 dB.

Two tone intermodulation measurements with a 20 MHz spacing were done (Fig. 5). At the 1 dB compression point the third order intermodulation suppression is 30 dBc.

The measured fundamental, the power added efficiency and the drain efficiency at 17.2 GHz agree with the harmonic balance simulation as shown in Fig. 6. The measured power added efficiency has a maximum of 20.1 % and the drain efficiency is 31.5 %.

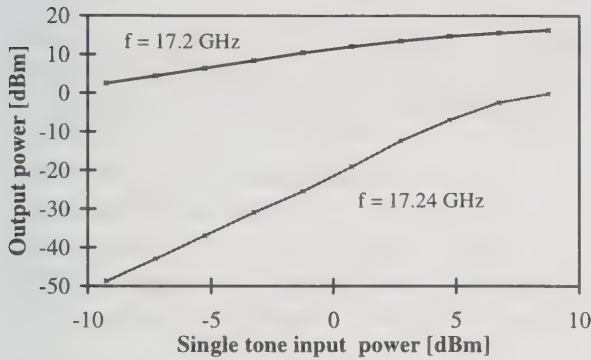


Fig. 5: Measured fundamental and third order intermodulation product (fundamental frequencies: 17.2 and 17.22 GHz)

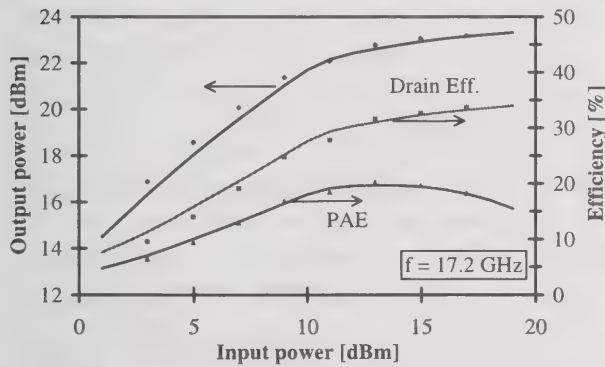


Fig. 6: Measured (dotted) and simulated (line) fundamental output power, drain efficiency and power added efficiency (PAE)

The effects of changes in the supply voltage, occurring in battery powered portable computers, are shown in Fig. 7. At 4.0 V supply of both transistors the saturated output power remains at 23.1 dB, but the gain decreases. At 2.5 V supply the gain as well as the saturated output power are

reduced. The amplifier has its maximum gain and output power at 3.3 V.

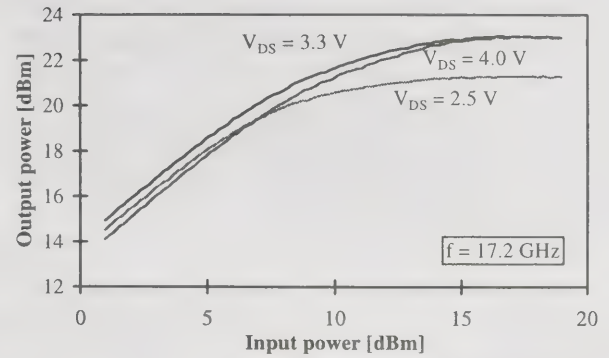


Fig. 7: Output power as a function of supply voltage

Small signal measurements (Fig. 8) show the spread of the samples. The better two samples have a gain variation less than 1 dB in the frequency range from 17 to 20.5 GHz. An additional feature of the designed power amplifier is a low output reflection coefficient of at least -10 dB over the frequency range (Fig. 9). The output match makes an application of the amplifier as driver stage for amplifiers with higher power outputs possible.

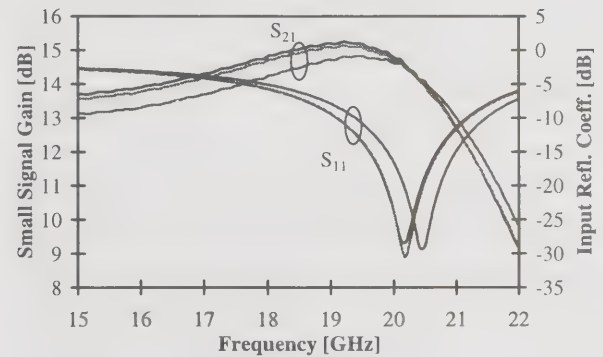


Fig. 8: Small signal gain and input reflection coefficient

The broadband power performance measured at 17.2, 18, 19 and 20 GHz is shown in Fig. 10. At 18 GHz the amplifier performance is better than at 17.2 GHz. The saturated output power slightly increases by 0.3 dB, and the output power at the

1 dB compression point is at 20.9 dBm. The efficiencies increase to PAE = 24.1 % and drain efficiency = 38.7 %. The efficiencies decrease for frequencies higher than 18 GHz. The output power and the power at the 1 dB compression point vary less than 2 dB over the frequency range. The third order intercept point, calculated for single tone output power (Fig. 5), is also shown in Fig. 10. It decreases from 29 dBm at 18 GHz to 27 dBm at 20 GHz.

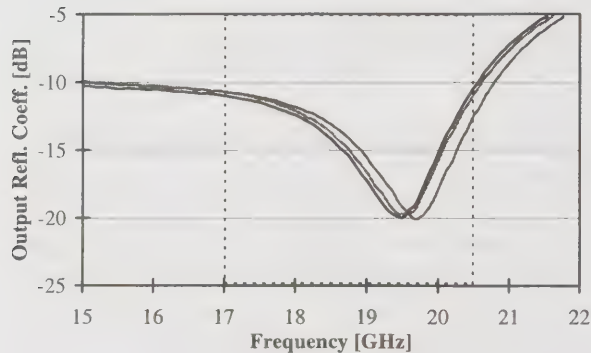


Fig. 9: Output reflection coefficient

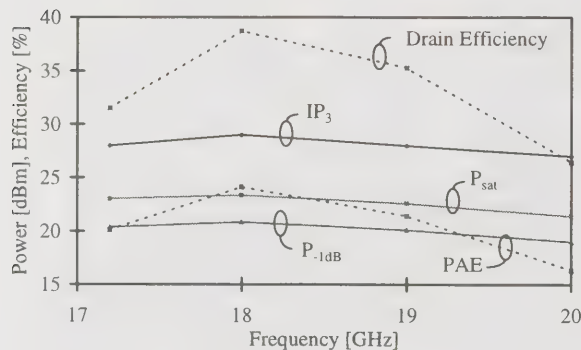


Fig. 10: Broadband performance of the amplifier

Conclusions

A HEMT amplifier with on-chip matching and bias networks was designed for wireless LAN with a commercial process. At 17.2 GHz an output power of 23.1 dBm and 14.1 dB power gain were measured. With a supply voltage of 3.3 V an efficiency of 31.5 % was measured at 17 GHz. Up to 20 GHz the amplifier has an output power higher than 20 dBm and a low output reflection coefficient.

Acknowledgment

This work has been performed in the framework of the project ACTS AC085 The Magic WAND, which is partly funded by the European Community and the Swiss BBW (Bundesamt für Bildung und Wissenschaft). The authors would like to acknowledge the contributions of their colleagues from Nokia Mobile Phones, Tampere University of Technology, Technical Research Centre of Finland, Ascom Tech AG, Lucent Technologies WCND, University of Lancaster, Robert BOSCH GmbH, University of Ulm, Compagnie IBM France, IBM Zürich Research Laboratory, ETH Zürich, INTRACOM Hellenic Telecommunications and University of Athens.

References

- [1] R. O. LaMaire, A. Krishna, P. Bhagwat, "Wireless LANs and mobile networking: Standards and future directions", IEEE Communications Magazine, August 1996, pp. 86 - 94
- [2] E. Ayanoglu, K. Y. Eng, M. J. Karol, "Wireless ATM: Limits, challenges and proposals", IEEE Personal Communications, August 1996, pp. 18 - 34
- [3] K. Pahlavan, A. Zahedi, P. Krishnamurthy, "Wideband local access: Wireless LAN and Wireless ATM", IEEE Communications Magazine, November 1997, pp. 34 - 40
- [3] J. Mikkonen, C. Corrado, C. Evci, M. Prögler, "Emerging wireless broadband networks", IEEE Communications Magazine, Feb. 1998, pp. 112 - 117
- [4] J. Aldis, E. Busking, T. Kleinje, R. Kopmeiers, R. van Nee, R. Mann-Pelz, T. Mark, "Magic into reality, building the WAND Modem", Proc. of ACTS Mobile Summit 1997, Aalborg, Denmark
- [5] S. Sumei, "Broadband measurements of indoor Wireless LAN at 18 GHz using directive antennas", Proc. of 1996 Intern. Conf. on Communication Technology, Beijing, China, May 1996, pp. 474- 477

A 85 Mbps Low-Power DQPSK MODEM-LSI for Advanced Wireless Access

Akihiro YAMAGISHI*, Masahiro MURAGUCHI** and Tsuneo TSUKAHARA*
NTT System Electronics Laboratories **NTT Wireless Systems Laboratories

3-1 Morinosato Wakamiya Atsugi-shi, Kanagawa, 243-0198, Japan
Phone: +81 462 40 2198, Fax: +81 462 40 4219, E-mail: aki@aecl.ntt.co.jp

Abstract

A new differential quadrature phase shift keying MODEM-LSI uses differential mapping for modulation and baseband delay detector logic for demodulation. The LSI achieves a maximum bit rate of over 85 Mbps and low power dissipation of 67 mW at 80 Mbps with a 3-V power supply. We built an intermediate-frequency (IF)-part test circuit composed of the new MODEM-LSI, and did a transmission test at IF return. The C/N degradation of this circuit is about 1.7 dB at 75 Mbps.

Key word: MODEM, DQPSK, differential detector, CMOS, wireless systems

Introduction

Wireless systems will provide wireless extensions of broadband multimedia networks. In multimedia service environments, users will send or receive various type of data, such as text and motion-video, to or from various sites. The asynchronous transfer mode (ATM) is well suited for these environments since it provides flexible bit rates and connections and supports various services. These capabilities will require new wireless systems, such as the ATM wireless access system (AWA) proposed by NTT[2]. In AWA, a user's wireless terminal, which consists of an ATM interface baseband digital signal processing portion, RF/IF portion, antenna, and battery, is attached to a notebook PC. The AWA can use the quasi-millimeter or millimeter wave band and has a modulation speed of more than several tens of Mbps. To encourage the spread of services like AWA, we must reduce the size of wireless terminal and make them more power efficient. The LSIs are the keys to the creation of compact terminals. At WCC'97, we introduced a 2-GHz Si-bipolar chip-set consisting of a logarithmic/limiting amplifier and a quadrature modulator and demodulator[1]. In this paper, a new differential quadrature phase shift

keying (DQPSK) MODEM-LSI for compact wireless terminals is proposed and demonstrated. The LSI achieves modulation and demodulation bit rates of more than 80 Mbps yet consumes very little power, and the chip is small. We got good performance in an intermediate frequency (IF) return transmission test using the MODEM-LSI and Si-bipolar chip-set.

MODEM-LSI

The new LSI includes a baseband logic circuit as a modulator and demodulator for differential quadrature modulation (DQPSK). Figure 1 is a block diagram of the LSI. The baseband logic for the modulator is composed of an input data mode selector, mapping circuit, and voltage-to-current converter (V-I converter). The mode selector selects serial or parallel and gray or natural input data types. The mapping circuit is a summing circuit, i.e., a two-bit accumulator, for delay detection. The V-I converter changes digital signals to analog current signals for the interface with the quadrature modulator. The base-band logic for the demodulator is a delay detector logic circuit composing a phase detector, a delay element, and a discriminator. The phase detector

generates 8-bit phase data from 8-bit I- and Q-channel data from the quadrature demodulator. The phase detector uses a ROM table. However, the ROM occupies a large amount of space and dissipates a lot of power, and the LSI's maximum bit rate is decided by the ROM throughput rate. Hence, it is very important to reduce the ROM storage capacity. In this LSI, the phase detector reduces the ROM storage capacity using a technique based on phase plane symmetry. Figure 2 shows an example of the phase plane. If all phase data, from 0 to 2π radians, were generated by the ROM, the ROM storage capacity would be 512 kbits, which is equal to 64k-word input and 8-bit output. However, in this phase detector, phase data from 0 to $\pi/2$ radians is generated by the ROM and phase data from $\pi/2$ to 2π radians is generated logically by the circuit. Figure 3 is a block diagram of the phase detector. In this circuit, ROM storage capacity is only 96 kbits, which is equal to 16k-word input and 6-bit output. Here, '1's comp' is a one's complementary calculation circuit composed of exclusive-OR gates. The delay element is a D-flip flop, and the discriminator is composed of an 8-bit adder and a few gates. Figure 6 is a photograph of the MODEM-LSI. The ROM would occupy large area in this LSI, were it not for the ROM reduction technique. The chip is 4 mm square. The process technology is $0.5\mu\text{m}$ CMOS process technology.

Experimental result

Figure 5 shows C/N versus bit error rate (BER) characteristics of the MODEM-LSI. The C/N degradation is 0.5 dB in the modulator and 0.6 dB in the demodulator at the BER of 10^{-5} . The power dissipation versus bit-rate characteristics are shown in Fig. 6. The power dissipation is 67 mW at 80 Mbps and 10 mW at 0 Mbps. This dc current is dissipated by the ROM. The MODEM-LSI performance is summarized in Table 1. We made an IF part test circuit composed of the new MODEM-LSI and Si-bipolar chip set proposed in WCC'97 and tested transmission at IF return. Figure 7 is a block diagram of the test circuit. The quadrature modulator, quadrature demodulator and logarithmic/limiting amplifier use the Si-bipolar chip set.

The analog-to-digital converters are commercially available LSI's, and the roll-off filters are SAW filters. The transmission bit rate is 75 Mbps I- and Q-channel 37.5 Mbps each. The intermediate frequency is 150 MHz in this circuit. Figure 8 shows the C/N versus BER characteristics of the test circuit. The C/N degradation is about 1.7 dB at the BER of 10^{-5} . This result confirms this component is applicable to advanced wireless access systems.

Conclusion

A new DQPSK MODEM-LSI for advanced wireless access has been presented. Power dissipation is only 67 mW at 80 Mbps, which is only 10% the power dissipation of the same functional block made by conventional parts. The maximum bit rate is 85 Mbps. The LSI is useful for broadband wireless communication terminals. The LSI, mounted on the 80 pin QFP package, is 12 mm square, which makes this a modem less than a quarter the size of those composed of conventional components. Experimental results for an IF block composed of new LSI and Si-bipolar chip sets show this circuit is applicable for advanced wireless access systems. The MODEM-LSI and IF Si-chip set can support the development of compact wireless terminals.

Acknowledgments

The authors would like to acknowledge T. Murase and J. Yamada for their guidance and encouragement, Y. Urabe and T. Arimoto for simulation and chip layout and T. Okamoto for measurements.

References

- [1] T. Tsukahara, K. Shiojima and M. Ishikawa, "A GHz-Band IF Si Chip-Set and Dielectric Root-Nyquist Filter for Broadband Wireless System", Proc. of WCC'97, pp. 32-37, Aug., 1997.
- [2] M. Umehira et al., "An ATM Wireless Access System for Tetherless Multimedia Services", ICUPC'95, PP. 858-862, 1995.

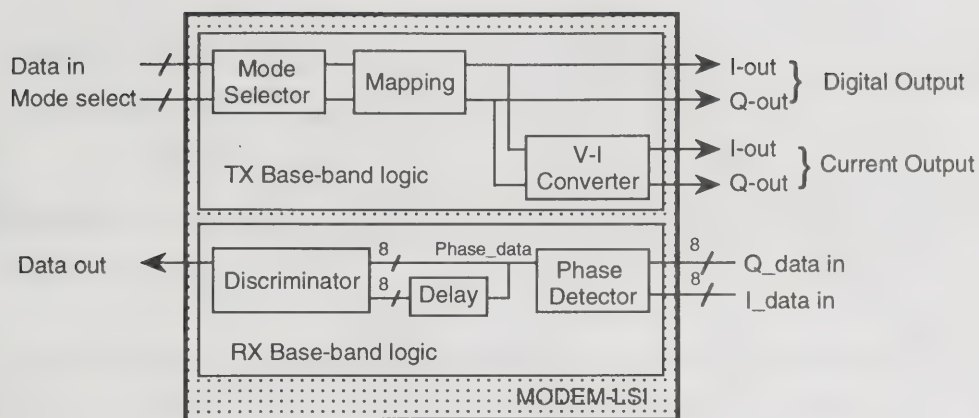


Fig.1 Block Diagram of the MODEM-LSI

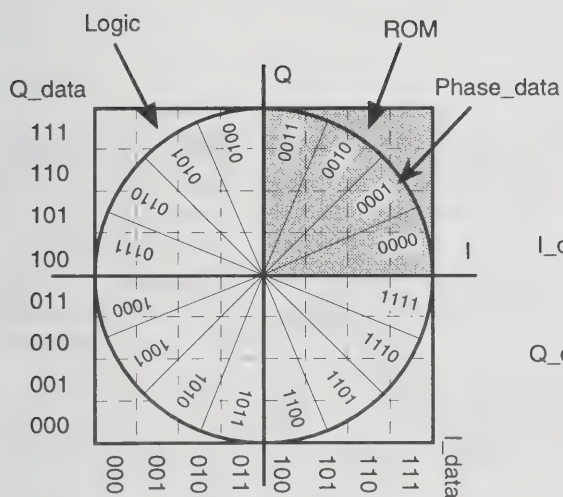


Fig. 2 Example of the phase plane

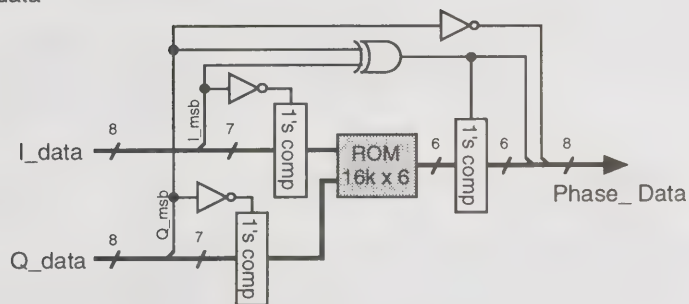


Fig. 3 Phase detector

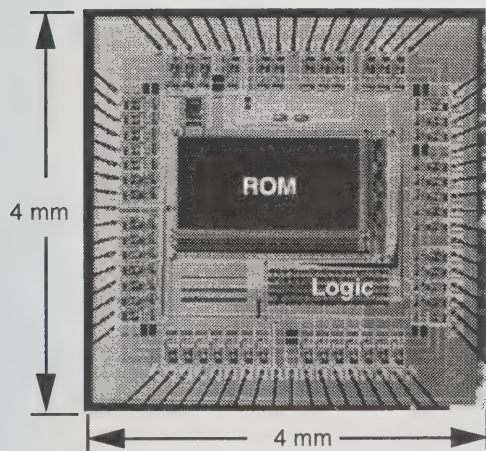


Fig. 4 Photograph of the MODEM-LSI chip

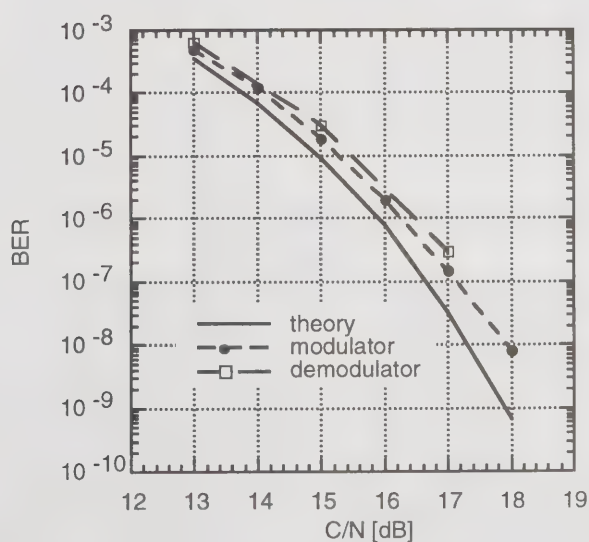


Fig. 5 C/N vs. BER

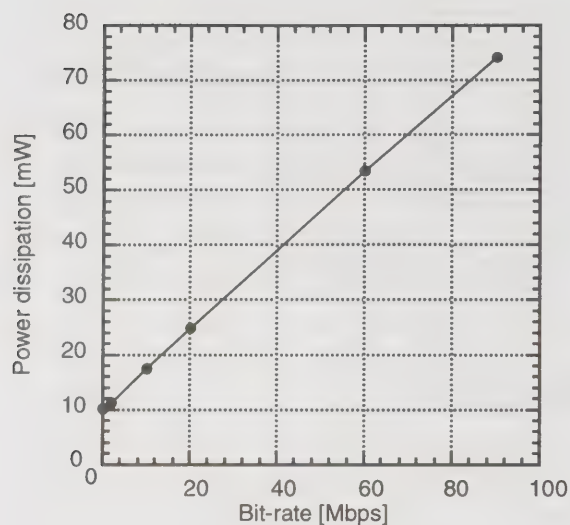


Fig. 6 Power dissipation

Table 1 MODEM-LSI Performance

Power Supply	3/3.3 V
Power Dissipation @80Mbps	67 mW
Maximum Bit Rate	>85 Mbps
C/N degradation	
TX	0.4 ~ 0.5 dB
RX	0.4 ~ 0.6 dB
Chip Size	4 mm ²
Process Technology	0.5 μ m-CMOS
Package Type	80 pin-QFP
Package Size	12 mm ²

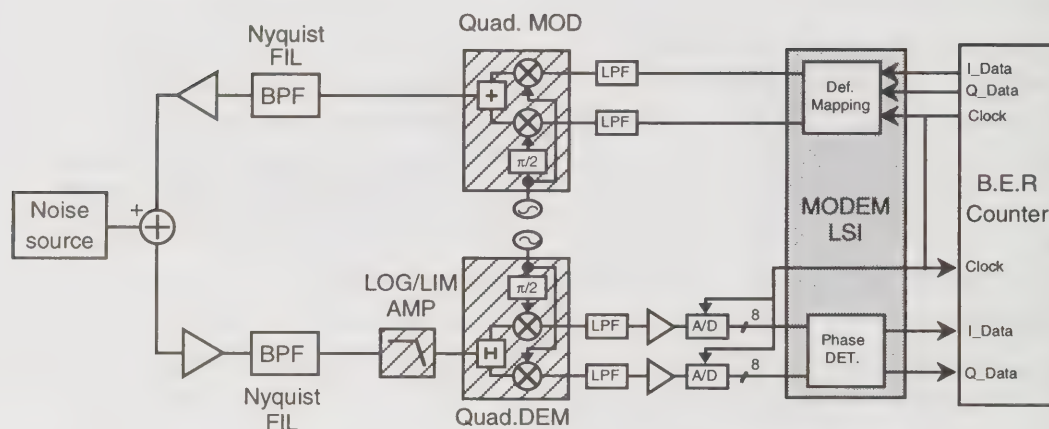


Fig. 7 Block diagram of the test circuit

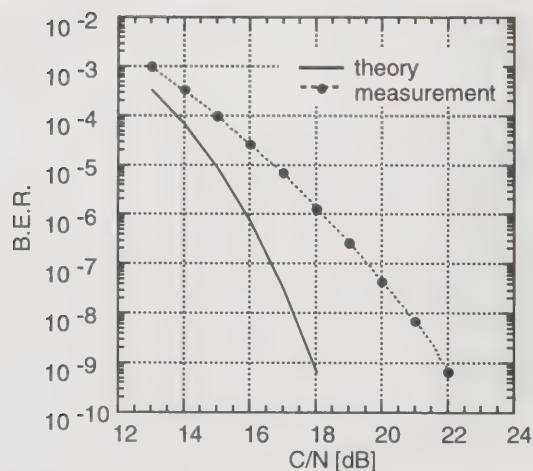


Fig. 8 C/N vs. BER of the test circuit

A Cost-Effective Approach to a Short-Range, High-Speed Radio Design in the U-NII 5.x GHz Band

Mehmet Soyuer, Herschel A. Ainspan, Joachim N. Burghartz, Jean-Olivier Plouchart, Brian P. Gaucher, Troy J. Beukema, Frank J. Canora, Erik Pilmanis and Modest M. Oprysko
soyuer@us.ibm.com, fax: 914-945-1974

IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598

Abstract

High-speed radio design research activities undertaken by the Communications Technology Department at IBM's T.J. Watson Research Center are described. The ultimate goal is to implement a cost-effective, short-range, high-speed radio in the 5-GHz U-NII band and leverage IBM's ultra-high bandwidth SiGe BiCMOS technology.

1.0 Introduction

On January 9, 1997 the FCC allocated 300 MHz of spectrum in the 5 GHz region specifically dedicated for short-range, high-speed wireless digital communications using unlicensed devices. The band is termed the Unlicensed-National Information Infrastructure (U-NII). Major corporations in the computer industry are participating in workgroups to define standards for RF Home Networking to exploit the bandwidth available in the U-NII band. However, at this time, there is no low-cost means for developing products in the U-NII band based on existing cellular, cordless, satellite, etc. technologies. IBM is in the position to develop and offer a solution based on its SiGe BiCMOS technology and exploit that technology through its business units. Our goal is to demonstrate that IBM has the base technology required for a cost-effective, short-range, high-speed digital communications in the U-NII band and to take the first steps in developing a design methodology, framework, and macros for the 5.x-GHz band including circuits and devices comprising the RF front-end, data converters, and baseband processor. The know-how gained in this work will also be directly

applicable to the band of frequencies (5.15-5.30 GHz) used in Hiperlan, a wireless LAN standard recently adopted by European Telecommunications Standards Institute (ETSI).

2.0 5-GHz Microwave Transceiver

The new U-NII band offers an unprecedented opportunity to develop new high-speed communication devices. The FCC has purposely limited the number of regulations developers must comply with. Maximum power is limited to 50 mW for the 5.15-5.25 GHz band (indoor environment) and to 250 mW for the 5.25-5.35 GHz band (outdoor environment). Based on these FCC rules for the U-NII band and our experience with the 5-GHz SiGe test circuits and DSP algorithms, we have arrived at a preliminary set of specifications for a 5-GHz radio which is outlined in Table 1.

TABLE 1. Single-chip 5-GHz transceiver target specifications

Parameter	Target Specification
Output Power	50-100 mW
Frequency Band	5.15-5.35 GHz (FCC U-NII)
Channel Spacing	25 MHz
# of Channels	7
Receiver Sensitivity	<-85 dBm
Data Rate	Variable 5 to 15 Mbps
Range	>75 ft

The ultimate goal for this project is to implement a low cost transceiver. This requires the judicious choice of an architecture which eliminates most of

the high-frequency filters such as direct conversion or single-IF conversion and a high-level of integration to reduce parts count and packaging cost. Direct conversion architectures have been successfully used for pagers and by Alcatel for its GSM phones. However, one must address the challenging issues such as DC offset and VCO leakage to the antenna. This problem is partially addressed in the macros themselves. Integration poses the challenges of noise, intermodulation distortion and self-jamming. Therefore, one of the short-term goals of this project is to study the isolation and cross-talk issues at microwave frequencies and compare different radio architectures.

A generic radio block diagram is shown in Figure 1. The building blocks that are being developed for the microwave transceiver section are LNAs, PAs, VCOs, mixers, phase shifters and frequency synthesizers. Some of our first-pass hardware results in a SiGe bipolar technology are summarized in [1]-[5]. In this paper, measured results for new LNA, mixer and VCO circuits implemented in a SiGe BiCMOS technology are presented.

2.1 Low-Noise Amplifier (LNA)

A 5.5-GHz LNA with switchable gain and temperature compensation is implemented in our 200-mm SiGe BiCMOS technology. This three-level-metal (Al) technology, with a substrate resistivity of 10-20 Ω -cm, features a SiGe HBT with a cutoff and maximum oscillation frequency in excess of 45 and 60 GHz, respectively, nMOS and pMOS devices with a nominal L_{eff} of 0.35 μ m, polysilicon and diffused resistors, high-density MOS and high-Q metal-insulator-metal (MIM) capacitors, varactor diodes, and inductors [6].

The LNA is comprised of two sections: the amplifier core and the proportional-to-absolute-temperature (PTAT) current reference. The amplifier core schematic is shown in Figure 2. This circuit consists of two parallel cascode amplifier stages. A pMOS switch is used to select either the high or low-gain mode. On-chip inductors have a quality factor of about 8 at 5.5 GHz.

The measurements are done at wafer level. The measured LNA gain and 50- Ω noise figure are 14.1 dB and 2.4 dB, respectively, at 25°C. The variation in gain and noise figure is less than 0.5 dB from 0 to 100°C. The input and output return loss in the high-gain mode are about 6 and 9 dB, respectively, at

25°C. The input third-order intermodulation intercept (IIP3) is +1.4 dBm. In the low-gain (or attenuation) mode, the circuit provides about 10.4 dB loss. The circuit consumes 4.7 mA total current from a 2.5-V supply. The LNA figure of merit, defined as the ratio of peak gain to the product of 50- Ω noise figure and DC power dissipation, is 0.5 mW^{-1} for this amplifier. This number is better than previously published figure-of-merit values for GaAs LNA circuits, e.g. see the tables and references in [1]-[2].

2.2 Mixer

A low-noise double-balanced active mixer, or Gilbert cell, is designed using the same SiGe BiCMOS technology described above. A schematic of the mixer is shown in Figure 3. Because of its low local-oscillator (LO) leakage, high conversion gain and differential architecture, this type of mixer is suitable for highly-integrated microwave receivers. The sizes of all differential pair transistors are optimized to get simultaneous input impedance and optimum noise matching. The usual transistor current source is replaced by an inductor, providing a DC path for the bias current and a high impedance at 5 GHz at the same time.

The mixer is measured on wafer between 4.5 and 6 GHz with an IF frequency of 100 MHz. The circuit exhibits a differential power conversion gain of 16.4 dB, an input 1-dB compression point of -20.7 dBm, an IIP3 of -11.1 dBm, and a double-side band noise figure of 6.6 dB at 5.2 GHz. A return loss better than 8 dB is measured from 5 to 6 GHz. The mixer core draws 10 mA from 3-V supply. Overall, a single-side band noise figure less than 10 dB is achieved thanks to the low-noise/high-gain characteristics of SiGe HBTs and by carefully sizing the transistors and passives.

2.3 Voltage-Controlled Oscillator (VCO)

VCOs are key building blocks which present many challenges to the designers of wireless ICs and systems. Full integration of the resonator tank circuit, low phase noise, low power, wide tuning range and a differential architecture are highly desirable for the integration of the VCO with other functions into a low-cost transceiver. For a given technology, there is a trade-off among frequency of oscillation, phase noise and power consumption. Furthermore, it is usually difficult to obtain low phase noise and high tuning

range at the same time. However, this paper presents a fully-integrated 5.6-GHz VCO which meets all of these challenges by exploiting the low-noise and high-gain characteristics of IBM's SiGe HBT technology.

Figure 4 shows the simplified schematic of the differential VCO. A differential pair combined with MIM capacitors in the positive feedback path forms the negative-resistance cell. Two identical, fully-monolithic resonators terminate the outputs of the differential-pair gain stage.

The VCO circuit dissipates a total of 66 mW from a 3-V supply. One third of the total current is consumed in the VCO core and two thirds in the buffer stages. The VCO tuning range is 15% for a varactor voltage variation between 0 and 3 V. The phase noise varies from -90 to -94 dBc/Hz at 100-kHz offset and from -110 to -116 dBc/Hz at 1-MHz offset across the full tuning range. Less than 170-MHz center-frequency drift is measured for a temperature variation from -50 to 150°C. Across the full temperature and supply voltage range (2.7 to 3.3 V), the center frequency variation is 3.3%, which is well below the tuning range as desired. The single-ended output power varies from 0 to -5 dBm over the same measurement range.

3.0 Baseband DSP

A modulation scheme has been developed which provides increased spectral efficiency over traditional Direct Sequence-Spread Spectrum (DS-SS) technology while maintaining acceptable system performance in an indoor multipath fading channel. The technique developed is based on a set of multipath resistant waveform codes (MRWC) which simultaneously provide signal spreading and data modulation. The result is a robust coding technique based on a simple modulation (e.g. BPSK/QPSK) which can be inexpensively implemented. It is estimated that the baseband processor can be realized in a relatively small ASIC of approximately 50k gates, since a complicated equalization scheme is not required in the demodulator.

The MRWC technology can provide a 2.5x to 5x data rate increase over the existing 1/2 Mb/s IEEE802.11 DS-SS standard while using the same spectral bandwidth and exhibiting enough processing gain to provide reliable operation in an indoor propagation environment. The increase in data rate comes at a

cost of a slightly more complicated correlator bank based receiver and reduced coverage range. The reduction in coverage range can be compensated through a diversity antenna reception system and/or a moderate increase in transmitter power, depending on the desired data rate. Variants of the design have also been developed to provide a 10 Mb/s data rate in a wider channel bandwidth, while avoiding the cost and complexity of a diversity antenna system, and lowering the amount of transmitter power required for robust system operation.

The signal processing algorithms which implement the waveform coding modulation have been prototyped on a pseudo real-time DSP VME system using a fast dual ADC to digitize the radio receiver I/Q channels, a 64 Msample memory buffer, and a Quad C40 DSP engine. Using this system in conjunction with standard DS-SS radio front-ends, we have been able to demonstrate laboratory data transmission rates from 5 to 38 Mb/s, with emphasis on providing robust 5 and 10 Mb/s data rates in the 2.4 GHz and 5 GHz ISM and NII radio bands.

4.0 References

- [1] H. Ainspan et al., "A 6.25-GHz low DC power low-noise amplifier in SiGe," Proc. of the IEEE Custom Integrated Circuits Conference, pp. 177-180, May 1997.
- [2] M. Soyuer et al., "A 5.8-GHz 1-V low-noise amplifier in SiGe bipolar technology," Dig. of Tech. Papers, IEEE Radio Frequency IC Symposium, pp. 19-22, June 1997.
- [3] M. Soyuer et al., "An 11-GHz 3-V SiGe voltage-controlled oscillator with integrated resonator," IEEE Jour. Solid-State Circuits, vol. 32, pp. 1451-1454, Sept. 1997.
- [4] J. Burghartz et al., "Integrated RF components in a SiGe bipolar technology," IEEE Jour. Solid-State Circuits, vol. 32, pp. 1440-1445, Sept. 1997.
- [5] J. Burghartz et al., "RF circuit design aspects of spiral inductors on silicon," Dig. of Tech. Papers, IEEE International Solid-State Circuits Conf., paper 16.1, pp. 246-247, Feb. 1998.
- [6] D. Ahlgren, et al., "A SiGe HBT BiCMOS Technology for Mixed-Signal RF Applications," Proc. of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting, pp. 195-198, Sept. 1997.

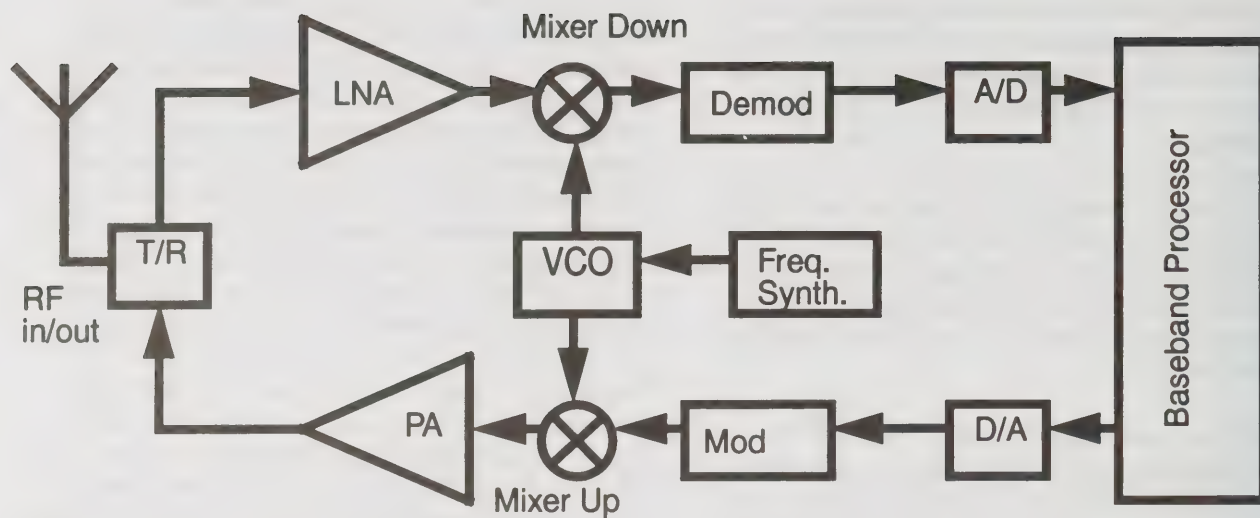


FIGURE 1. Generic radio block diagram

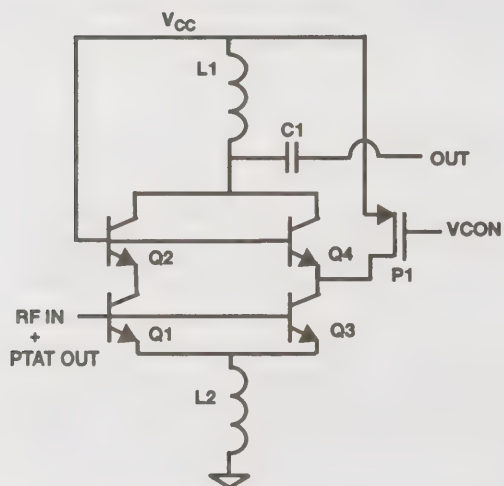


FIGURE 2. Amplifier core schematic

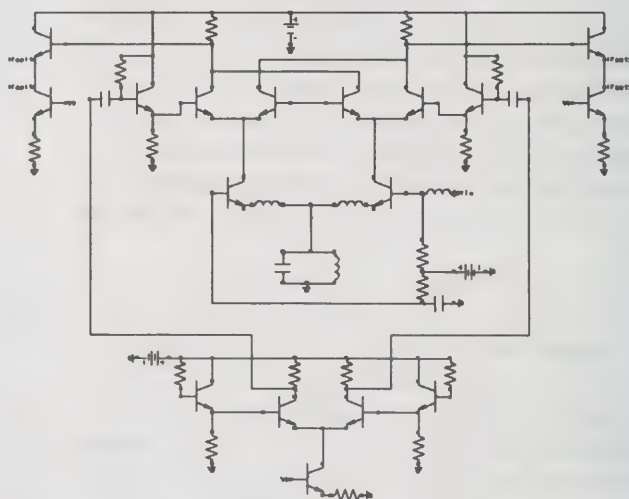


FIGURE 3. Mixer schematic

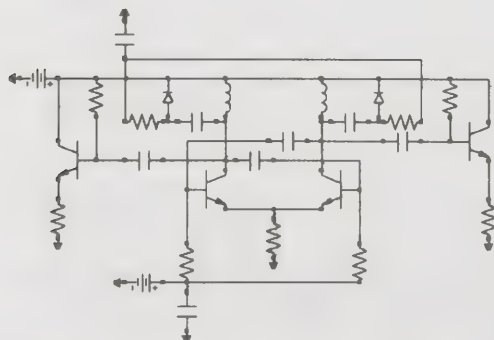


FIGURE 4. VCO schematic

A high-efficiency RF transmitter using VCO-derived synthesis: CALLUM

David J. Jennings and Joseph P. McGeehan

Centre for Communications Research, University of Bristol, Queens Building,

University Walk, Bristol BS8 1TR, United Kingdom.

Tel: 44 117 9287740 FAX: 44 117 954 5207

Email: david.jennings@bristol.ac.uk

Abstract

The combined analogue locked-loop universal modulator (CALLUM) is an RF transmitter topology which produces a linear output through the use of non-linear but highly efficient RF power amplifiers. This is achieved through careful phasing of two constant-envelope vectors, each derived from voltage-controlled oscillators. The system up-converts and amplifies a baseband signal within a closed-loop feedback scheme. The optimal CALLUM system requires complex baseband processing, but a simplified version, known as CALLUM2 can be implemented using simple analogue circuitry. This paper provides some insight into the behaviour of such systems and gives results from an experimental CALLUM2 system in response to a modulating signal compatible with that used in the TETRA standard.

Introduction

The ever-increasing use of mobile communications equipment has necessitated the use of spectrally efficient modulation schemes such as $\pi/4$ -QPSK. Such schemes require a linear transmitter to preserve the trajectory of the complex signal, thereby confining the transmitted power to a well defined bandwidth and not causing interference in adjacent channels. Whilst reasonable levels of linearity can be achieved from class A amplifiers, such arrangements are not power efficient and therefore are unsuitable for use in hand-portable equipment. Existing solutions for producing linear, power efficient transmitters involve using an amplifier stage of increased efficiency such as class C and then attempting to linearize the resulting output. Several linearization techniques exist, the most common of these being Cartesian loop, feedforward and predistortion. A more radical approach to producing a high efficiency, linear transmitter is to use the combined analogue locked-loop universal modulator (CALLUM). Unlike other systems, CALLUM offers the potential of ultra-high power efficiency nearing 100%. Clearly, such a transmitter would find extensive use in hand-portable (i.e. battery-operated) systems.

The CALLUM system produces a linear, high-level RF output signal in the same manner as the LINC (linear amplification with non-linear components) method proposed by Cox [1]. An arbitrary input signal varying in phase and amplitude is decomposed into two constant-envelope signals (fig. 1) which can each be amplified by grossly non-linear, but highly efficient amplifiers and

then recombined at high level to synthesize the required output. The constant-envelope signals are derived from two voltage-controlled oscillators (VCOs). The original LINC system operated in an open-loop configuration and was sensitive to small gain and phase imbalances in the two component arms. The CALLUM system has a closed-loop topology and can continually correct for such imbalances. It is a complete transmitter scheme which both upconverts and linearly amplifies a baseband signal represented in Cartesian form, within a closed-loop system. Due to the closed-loop, high gain nature of this arrangement, the system is inherently narrowband. The bandwidth of the system is defined by the dominant poles and time delay around the loop which will eventually cause instability at higher loop gain levels.

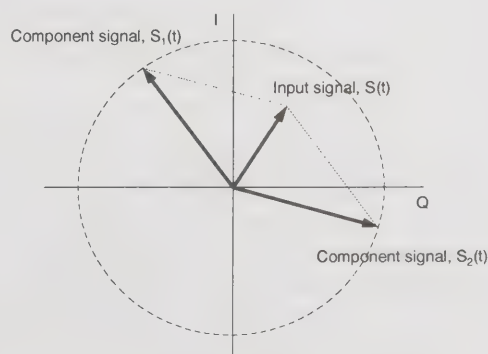


Figure 1: Decomposition of the input signal into two constant-envelope component signals.

System Description

In the generic form of CALLUM proposed by Bateman [2] the two VCOs are driven with the error signal generated by comparing the input and downconverted Cartesian components (fig. 2). This system is not stable for all input phases, however stability can be guaranteed if a sign-switching matrix is employed [3]. Unfortunately, this modified version suffers from discontinuities when the input signal changes from one quadrant of the complex plane to another.

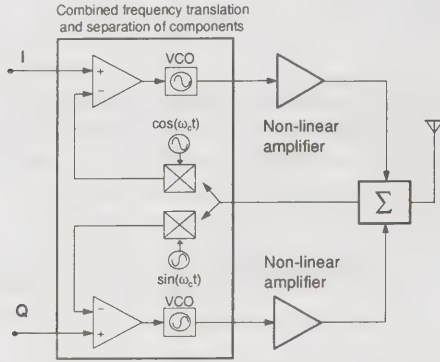


Figure 2: General CALLUM topology

A much improved version of CALLUM was subsequently proposed [4] which is stable for all input phases and does not suffer from discontinuities. The resulting system, known as CALLUM1, performs continual phase rotation and amplitude normalisation for optimal dynamic performance. To implement CALLUM1, the VCOs must be excited with control signals derived from functions of the baseband input signal and error signal vectors (plus known system constants):

$$v_1 = k\Delta x \left[\frac{-x}{cr\sqrt{4R^2 - r^2}} - \frac{y}{cr^2} \right] + k\Delta y \left[\frac{x}{cr^2} - \frac{y}{cr\sqrt{4R^2 - r^2}} \right]$$

$$v_2 = k\Delta x \left[\frac{x}{cr\sqrt{4R^2 - r^2}} - \frac{y}{cr^2} \right] + k\Delta y \left[\frac{x}{cr^2} + \frac{y}{cr\sqrt{4R^2 - r^2}} \right]$$

Eq. (1)

where x and y are the Cartesian components of the input signal, Δx and Δy are the corresponding error signals, c is the VCO sensitivity, k is the loop gain, r is the input signal envelope and R is the magnitude of the two VCO vectors. Implementing these equations is not easy; a practical method of generating v_1 and v_2 might be to employ a DSP. This would be feasible because the bracketed terms of (1) can be generated outside of the feedback loop, thus minimizing the effect of time delay on loop stability. It should be borne in mind however

that although the baseband input signal would typically be of the order of 20kHz, the non-linear nature of (1) can result in terms which have a bandwidth many times this figure. For a $\pi/4$ -DQPSK signal a ten-fold expansion can be expected. The closer to the complex origin that the input signal trajectory passes, the greater the bandwidth expansion.

Fortunately, CALLUM1 can be simplified and the need to use a DSP removed completely. If the denominators of equation (1) are ignored, then a version called CALLUM2 results. This is also stable, but unlike CALLUM1, its tracking performance degrades as the envelope of the input signal reduces. This is because loop gain (which directly affects tracking performance) is effectively reduced when the envelope becomes small (less than unity). When considering small step changes in the phase (or amplitude) of the input signal, the system can be considered first order and has a time constant equal to $1/k$. In terms of linear control theory, the system may be modelled as an integrator block with a unity-gain feedback path (fig. 3) when considering either phase or amplitude response. The resulting closed-loop transfer function is therefore:

$$G_{CL}(s) = \frac{k}{s + k} \quad (\text{Eq. 2})$$

which gives rise to the time constant of $1/k$. In the case of CALLUM2, the k in fig. 3 is replaced by kr^2 .

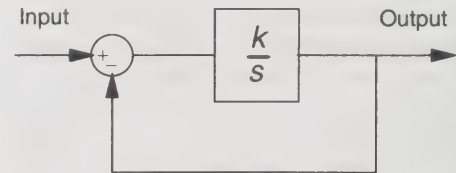


Figure 3: Linear model of CALLUM1 in response to a step change in phase

Tracking performance at low envelope levels is increased by the use of more loop gain, but at higher envelope levels the increased loop gain is likely to cause instability. A compromise between expected minimum/maximum envelope levels and loop gain has to be reached to obtain the best performance. In the case of CALLUM1, tracking performance is the same for any input envelope level (so long as the envelope level stays within the *synthesizable range* of $2R$ i.e. when the two components sum together in-phase. This is a fundamental condition for any LINC system). The level of loop-gain applied to the system also defines the maximum frequency component to which the system can lock [4], for CALLUM1 this is given by:

$$f_{\max} = \pm \frac{k}{2\pi r} \sqrt{4R^2 - r^2} \quad (\text{Eq.3})$$

and for CALLUM2,

$$f_{\max} = \pm \frac{kr}{2\pi} \sqrt{4R^2 - r^2} \quad (\text{Eq.4})$$

In the case of CALLUM1, small envelope levels result in high values of f_{\max} , the opposite is true for CALLUM2. Although lock may be achieved for frequency components less than f_{\max} , a large tracking error may exist and hence poor output linearity.

Hardware Implementation of CALLUM2

The CALLUM2 system can be implemented entirely in the analogue domain, as the two VCO control signals can be generated easily using standard 'linear' components. The structure of CALLUM2 is shown in fig. 4. The system was designed to operate at 220-240 MHz, although the choice of centre frequency is arbitrary as this has no affect on the baseband signal processing design. The core signal processing block is the multiplying matrix which was implemented using four Analog Devices AD-835 wideband multipliers. The error signal is generated using Analog Devices AD-830 differential amplifiers, as are the terms $x+y$ and $x-y$ (the AD-830 can be configured for resistorless summation). The $x+y$ term can be easily inverted to generate $-x-y$ as the AD-835 has differential inputs. The RF section includes the VCOs, non-linear power amplifiers, power combiner and downconverter.

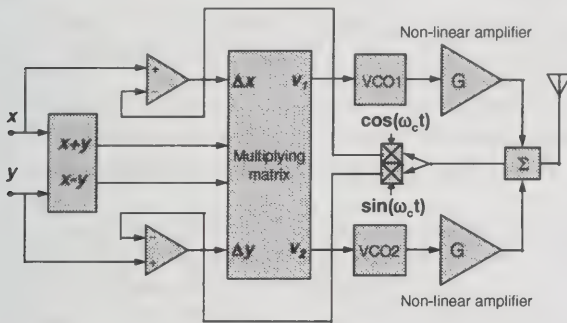


Figure 4: The structure of CALLUM2

Power modules: The PAs are Mitsubishi FM modules which operate at approximately 45% efficiency with a gain of 36dB (maximum output power = 7W).

Power combiner: Wilkinson type, constructed using two quarter wavelength, 75Ω co-axial cables lengths (ideally, this should be 70.7Ω) with a 100Ω load resis-

tor. The resistor must be adequate to dissipate all the power from the two PAs; when the synthesized envelope is very small the component signals sum whilst approaching anti-phase, the unwanted power must be dissipated in this load resistor. The power combining process is problematic, the power loss increases as the phase angle increases, ultimately dissipating all the power in the internal load when the signals are in anti-phase. A solution to this problem might be to use a summing-switching amplifier arrangement or by using a voltage summation method. It may be possible to use an RF operational amplifier if only modest output power is required.

Downconverter: This uses two Mini-Circuits TUF-1 mixers to effect direct quadrature downconversion. The LO is split into quadrature components using a Mini-Circuits PSCQ-2-250. The IF output (baseband in this case), is then amplified using operational amplifiers. These amplifiers have variable gain and a DC offset adjustment so that compensation can be made for any quadrature gain/phase imbalance.

VCO: This is the main element of the CALLUM system. At first, the VCOs were implemented using the DC-FM facility of two RF signal sources. The limited modulation port bandwidth of these sources (approximately 100 kHz) resulted in a low frequency pole in the feedback loop, causing instability at modest loop-gain levels. The onset of oscillation was signalled by an increase in the noise skirt at ± 90 kHz either side of the centre frequency. To improve the bandwidth of the system and to allow the use of more loop-gain, the signal sources were replaced by Z-COMM 930MHz VCOs. The output of the VCO is divided by a factor of 4 using a prescaler, this decreased the tuning sensitivity from 8 to 2 MHz/V and also improved phase noise performance. The sensitivity of the VCO directly affects loop gain; 2 MHz/V is equivalent to a loop-gain of 2,000,000. To set the free-running frequency and to allow a bipolar control voltage, the VCOs were biased by summing in a DC level. During set-up, this level is adjusted so as to have a free-running frequency which is the same as that of the intended LO. It is essential that the VCOs are DC-coupled for the system to operate. The tuning port bandwidth of these VCOs is in excess of 50 MHz and so shouldn't limit system stability. The output of the VCO is adequately buffered (to prevent undesired loading effects) by a series attenuator followed by a MMIC amplifier stage. The overall output power of the complete VCO blocks (including prescaler) is -6 dBm. The output of each PA is therefore approximately 1W.

A source of low frequency poles are the operational amplifier stages in the downconverter. Devices with high gain-bandwidth product were used so that the bandwidth

of the system could be made as large as possible. Loop gains twenty times greater could now be achieved, until oscillation at ± 3 MHz ensued.

The entire system is locked to the LO reference, therefore changing the transmit channel becomes a simple matter of changing the LO frequency to the new channel. The new channel must, however, be well within the linearization bandwidth of the system. For example, if a 1kHz tone was being transmitted, the tone would appear at $f_c + 1$ kHz. If the LO was changed to $f_c + 10$ kHz, then the system is then effectively attempting to track a tone of 11kHz which demands more loop gain for the same tracking performance. If the VCO free-running frequency were also changed to the new channel frequency then the system would revert back to tracking a 1kHz tone and have the associated benefit of better tracking performance.

Results from the hardware system show that linearity is good enough to achieve the transmission masks for most modulation formats including TETRA (trans-european trunked radio) so long as adequate stable loop gain can be provided. TETRA employs $\pi/4$ -DQPSK modulation at a data rate of 18kBaud with a filter alpha of 0.35. When excited by a TETRA modulation signal, the output spectrum shown in fig. 5 results. This figure also shows the spectrum of one of the two wideband, phase-modulated component signals. When these signals are summed, frequency components outside of the complex modulation bandwidth combine in anti-phase and cancel, leaving the desired linearly amplified output.

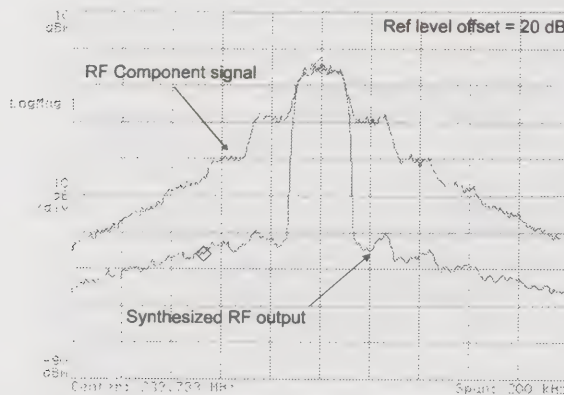


Figure 5: CALLUM2 output and component spectrum for a TETRA-type modulation input ($\pi/4$ -DQPSK, 18kHz, $\alpha=0.35$)

Conclusions

The CALLUM system provides a method by which highly efficient, switching-type non-linear RF power amplifiers can be used to generate a linear output. The potential of very high power efficiency from these sys-

tems makes them particularly suited to battery-operated, mobile applications. Overall system efficiency is however limited at present by power combining methods. Due to its closed-loop nature, CALLUM is particularly suited to narrowband systems such as TETRA or D-AMPS. The optimal CALLUM1 system requires non-linear calculations to be performed to obtain the ideal VCO control voltages. If these non-linear operations are ignored, a simplified system known as CALLUM2 results. This version can be implemented using simple analogue components, but suffers from poor tracking performance at low envelope levels. This is due to the dependence of loop gain on envelope level. If the envelope of the modulation signal does not vary over a wide range (as is the case with $\pi/4$ -DQPSK) then CALLUM2 can be particularly effective. The optimal CALLUM1 system has constant loop gain over the entire permissible envelope range and is therefore capable of tracking input signals whose envelope drops to very low values. If the envelope falls to zero, as is the case with a standard two-tone test then a discontinuity will occur, resulting in wideband distortion products. This applies to LINC systems in general, apart from those which use more than one component signal.

Further Work

To fulfil the potential of the CALLUM technology, a power efficient combining method must be sought, this is currently being researched. The optimal CALLUM1 system is currently under development by the authors and should give improved linearity, particularly with input signals with large envelope variations.

References

- [1] D. C. Cox, "Linear Amplification with Non-Linear Components", IEEE Trans. on Communications, vol. COM-22, pp. 1942-1945, December 1974.
- [2] A. Bateman, "The Combined Analogue Locked Loop Universal Modulator (CALLUM)", in Proceedings of the 42nd IEEE Vehicular Technology Conference, May 1992.
- [3] K. Y. Chan and A. Bateman, "Analytical and measured performance of the combined analogue locked loop universal modulator (CALLUM)", IEE Proceedings on Communications, vol. 142, no. 5, pp. 297-306, October 1995.
- [4] K. Y. Chan and A. Bateman, "Linear Modulators Based on RF Synthesis: Realization and Analysis", IEEE Trans. on Circuits and Systems: Fundamental theory and Applications, vol. 42, no.6, pp. 321-333, June 1995.

Bit Error Rate Analysis in Wireless Systems Modeling

George Chrisikos[†] and Christopher J. Clark

The Aerospace Corporation
P. O. Box 92957, M1-136 Los Angeles, CA 90009-2957
Tel.: 310-336-9244, Fax: 310-336-6225
E-mail: chrisiko@milly.usc.edu

ABSTRACT

This paper derives the bit error probability, or BER, of some binary modulation schemes as a function of the mean squared error between a degraded transmitted signal and the mismatched replica used at the receiver for correlation detection. This scenario is typically encountered in situations where nonlinear distortion from power amplification prevents the a priori knowledge of the signal waveform at the receiver, and a modeled waveform derived from time domain measurements is used.

I. INTRODUCTION

Specifications for wireless systems are primarily based on allocated channel bandwidth and minimum bit error probability, or bit error rate (BER). Tradeoffs in the system architecture, modulation format, and eventual hardware design are often selected based on a communications system model. Many of the current cellular systems have chosen non-constant envelope modulation schemes where power amplifier performance is critical. The primary focus in the literature has been on spectral regrowth [1] caused by the power amplifier nonlinearities.

Time domain envelope modeling has recently been developed to improve systems modeling and simulations using nonlinear elements [2]. The most popular modeling technique in system level design is based on the complex envelope of the time domain signals [3]. The recent development of accurate measurement techniques for the com-

The research described in this paper was supported by The Aerospace Corporation. The corresponding author [†] is also with the Communication Sciences Institute, Dept. of EE-Systems, University of Southern California, Los Angeles, CA 90089-2565.

plex envelope [4] has allowed for the development of more accurate models [5]. The improvement is specified in terms of the mean squared error (MSE) between the measured and modeled time domain complex envelope.

This paper focuses on the relationship between MSE and the bit error rate performance parameter. The results are useful in wireless systems design where the designer would like to estimate the accuracy of a specific model in predicting BER performance. The probability of error is calculated analytically as a function of the signal to noise ratio parameterized by the MSE. This allows one to predict the accuracy of a communication system model without resorting to time consuming computer simulations.

A simplifying assumption can be made on the nature of the waveform error. If the time domain error waveform is assumed to be Gaussian distributed, then the performance degradation results from the sum of noise powers due to the addition of two independent noise sources.

II. ANALYSIS

The bit error probability for digital communication systems is obtained by computing a statistical likelihood ratio as in [6]. The optimal decoder is shown to be a correlator or a filter matched to the transmitted waveform. When a mismatched filter is used due to the inability to properly construct the transmitted waveform, there is a resulting degradation to the BER. Under the assumption of additive white Gaussian thermal noise and no intersymbol interference, the following analysis computes the performance in terms of BER.

Let $s(t)$ denote the signal waveform which represents a *zero* or a *one* bit in a binary transmis-

sion system and let $\hat{s}(t)$ denote the mismatched waveform over a bit time T_b . With out loss of generality, assume $s(t)$ represents a *one*. The *mean squared error* (MSE) between waveforms is defined as

$$\text{MSE} = \int_0^{T_b} |s(t) - \hat{s}(t)|^2 dt. \quad (1)$$

Let the received signal at the demodulator be $r(t) = \hat{s}(t) + n(t)$ where $n(t)$ is an additive white Gaussian thermal noise waveform with variance $N_o/2$. The correlation detector computes the decision statistic $z = \int_0^{T_b} r(t)s(t) dt$ by correlating with the signal $s(t)$ and compares this value to a threshold (a threshold of zero is used for binary systems) to decide if a 0 or 1 bit was transmitted. The mean and variance of z is needed to compute the probability of error. The mean of z is denoted by μ_z and is given as

$$\mu_z = \int_0^{T_b} \hat{s}(t)s(t) dt \quad (2)$$

and the variance, σ_z^2 , is

$$\sigma_z^2 = \frac{N_o}{2} E_b \quad (3)$$

where E_b is the bit energy. The bit energy E_b and the energy in the mismatched waveform $E_{\hat{b}}$ are defined as

$$E_b = \int_0^{T_b} |s(t)|^2 dt \quad \text{and} \quad E_{\hat{b}} = \int_0^{T_b} |\hat{s}(t)|^2 dt. \quad (4)$$

The BER is derived by computing the statistics of the Gaussian random variable at the output of the detector and integrating the probability density function over the region corresponding to an error event.

It is easily shown that the probability of error for equally likely bits is given as

$$P_b(\mathcal{E}) = Q\left(\frac{\mu_z}{\sigma_z}\right) \quad (5)$$

where the function $Q(x)$ is the Gaussian tail integral defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du. \quad (6)$$

The waveform $\hat{s}(t)$ can be expressed as the sum of an error waveform and the desired waveform $s(t)$ as

$$\hat{s}(t) = s(t) + \epsilon(t). \quad (7)$$

In general, $\epsilon(t)$ is a deterministic error waveform but it can be modeled in terms of some statistical distribution. If the error waveform is random, then μ_z and σ_z^2 are actually random variables conditioned on ϵ .

Using the notation defined, the BER expression for a BPSK signal is given by

$$\begin{aligned} P_b(\mathcal{E}) &= Q\left(\sqrt{\frac{2E_b}{N_o}} \left[\frac{E_b + E_{\hat{b}} - \text{MSE}}{2E_b}\right]\right) \\ &= Q\left(\sqrt{\frac{2E_b}{N_o}} \left[1 + \frac{1}{E_b} \int_0^{T_b} s(t)\epsilon(t) dt\right]\right) \end{aligned} \quad (8)$$

This result shows that the signal to noise ratio E_b/N_o is degraded by the factor $[1 + \frac{1}{E_b} \int_0^{T_b} s(t)\epsilon(t) dt]^2$ which is zero dB if $\epsilon(t)$ is zero as expected.

A. Gaussian Distributed Error Waveform

The error waveform $\epsilon(t)$ between the measured and modeled response of wideband amplifiers has been shown to be Gaussian distributed in [7]. If the error waveform $\epsilon(t)$ is modeled as Gaussian noise with a zero mean and a variance of σ_ϵ^2 , then the expression given in (9) simplifies to

$$\begin{aligned} P_b(\mathcal{E}) &= Q\left(\sqrt{\left(\frac{2E_b}{N_o}\right)^{-1} + \left(\frac{E_b}{\sigma_\epsilon^2}\right)^{-1}}\right)^{-1} \\ &= Q\left(\sqrt{\frac{E_b}{N_o/2 + \sigma_\epsilon^2}}\right). \end{aligned} \quad (9)$$

The variance σ_ϵ^2 is an expression of the MSE and the result shows that this is equivalent to modeling the signal $\hat{s}(t)$ as the sum of the desired signal $s(t)$ and two independent noise sources $n(t)$ and $\epsilon(t)$; one due to thermal noise, the other due to the waveform modeling errors. The MSE is now a random variable whose mean, or average value, $\overline{\text{MSE}}$ is

$$\overline{\text{MSE}} = \sigma_\epsilon^2. \quad (10)$$

B. BER for Minimum Shift Keying

Another popular modulation used for wireless systems is *minimum shift keying* (MSK). This is a form of coherent FSK where the frequency deviation between the two waveforms used to transmit a 1 or 0 bit is one-half the bit rate. The resulting BER can be shown to be given by

$$P_{b,MSK}(\mathcal{E}) = Q\left(\sqrt{\frac{E_b}{N_o}} \left[\frac{E_b + E_b - \text{MSE}}{2E_b} \right]\right) \quad (11)$$

which is similar to (9) except for the factor of 2 in the numerator.

III. DISCUSSION

The expression in (9) is plotted in Fig. 1 which shows the bit error probability for BPSK modulation in terms of the percentage MSE, or precisely the average MSE normalized to the bit energy. For instance, the graph shows that at a BER of 10^{-6} , a 5% MSE gives about 3.5 dB degradation in E_b/N_o compared to the benchmark BPSK signal. Increasing levels of normalized MSE are shown in increments of 1%, which tend to move the BER curve to the right. At increasing levels of E_b/N_o , the separation between curves for increasing levels of MSE widens.

IV. CONCLUSION

This paper has presented the development of the BER in terms of the mismatch between measured and modeled waveforms. This is a useful parameter in the specification of wireless systems particularly in situations where time domain models are constructed to predict nonlinear distortion.

REFERENCES

- [1] J. Staudinger, "An examination of two measures of power amplifier linearity— Intermodulation distortion and channel spectral regrowth," in *Proc. 1996 Wireless Commun. Conf.*, 1996.
- [2] M. Heutmaker, J. R. Welch, and E. Wu, "Using digital modulation to measure and model RF amplifier distortion," in *Proc. 1996 Wireless Commun. Conf.*, 1996.
- [3] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. Plenum Press, New York, 1992.
- [4] A. A. Moulthrop, M. S. Muha, C. P. Silva, and C. J. Clark, "A new time-domain measurement technique

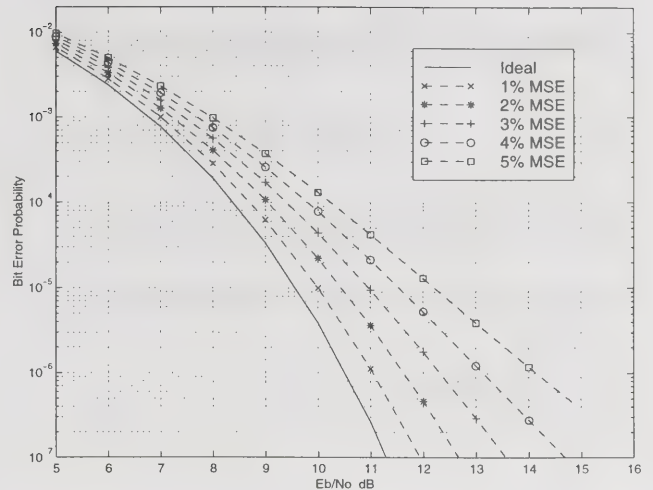


Fig. 1. Bit error probability for BPSK modulation as a function of the bit energy to noise ratio parameterized by the MSE.

for microwave devices," in *IEEE MTT-S International Symp.*, June 1998.

- [5] G. Chrisikos, C. J. Clark, A. A. Moulthrop, M. S. Muha, and C. P. Silva, "A nonlinear ARMA model for simulating power amplifiers," in *IEEE MTT-S International Symp.*, June 1998.
- [6] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*. Englewood Cliffs, New Jersey 07632: Prentice Hall, first ed., 1995.
- [7] C. J. Clark, G. Chrisikos, M. S. Muha, A. A. Moulthrop, and C. P. Silva, "Time-domain envelope measurement technique with application to wideband power amplifier modeling," submitted to *IEEE Transactions on MTT*, Dec. 1998.

Performance of an Adaptive Rate Modem Using Quasi-Analytic Simulation Techniques

Peter M. Hofstetter and Mark A. Wickert*

Department of Electrical and Computer Engineering

University of Colorado at Colorado Springs

P.O. Box 7150, Colorado Springs, CO, 80933-7150

719-262-3500, 719-262-3589 (fax), wickert@signal.uccs.edu

Abstract

Analysis of bit error rate (BER) of a π/M -MDPSK modem in an (additive white Gaussian noise) AWGN channel using quasi-analytic simulation techniques is considered in this paper. Implementation losses, most notable at $M = 32$ and 64 , which result from practical transmit and receive filtering, can be largely overcome by increased accuracy symbol timing recovery and a fixed equalization filter.

Introduction

The idea of an adaptive rate modem has received more and more attention in recent years due to the growing demand of transmitting high data rates over a mobile channel. The adaptive π/M M -ary differentially coherent phase-shift keying (MDPSK) modulation scheme is based on the North American TDMA standard for Data Communication, IS-130. The idea is to adapt the modulation index to the quality of the channel. The modulation scheme π/M -MDPSK, is a natural extension to the standard $\pi/4$ differentially coherent quadrature-shift keying (DQPSK) used in IS-130. In this paper we are primarily concerned with implementation losses associated with a DSP based modem. These implementation losses are primarily due to the intersymbol interference (ISI) introduced by the reconstruction and anti-aliasing filters, and imperfect symbol timing in the receiver decision sampler. The quasi-analytic (QA) simulation technique used in this paper is a combination of simulation and analysis. The main advantage is the reduction of simulation time compared to pure Monte-Carlo simulation. This is achieved by using a much shorter symbol sequence.

The π/M -MDPSK Modulation Scheme

The modem is based on the IS-130 standard which uses a $\pi/4$ -DQPSK modulation scheme. Since this system is designed for the worst case channel conditions, the user data rate is only 9.6 kbit/s. To increase the bit rate of the modem the modulation

scheme can be extended to the M -ary case. The M -ary source sequence $\{x_n\}$, $x_n \in \{\pm 1, \pm 3, \dots, \pm(M-3), \pm(M-1)\}$ is mapped to the differential phases $\{\pm\pi/M, \pm 3\pi/M, \dots, \pm(M-1)\pi/M\}$. The modulation index increases (up to 64 for this system) if the channel quality is good and decreases if the channel quality degrades.

Quasi-Analytic Simulation Techniques

A description of the QA simulation technique as applied to the estimation of symbol error probability, can be found in Jeruchim et. al. [JER92]. The overriding utility of the QA simulation technique is timesaving. The QA method uses both simulation and analysis to estimate the probability of error. The simulation is used to generate a noiseless waveform at the receiver. Given this waveform and assuming the noise is additive and has a known probability density function, the symbol error probability (SEP) can be calculated. In this work we collect noise free decision samples following differential detection and an interpolator. The differential phase error, ϕ_k , of each symbol decision sample and the envelope power deviation of the present and the previous symbol, β_k and β_{k-1} respectively, are obtained and averaged over the corresponding conditional SEP.

To obtain the symbol error probability we start with the $\pi/4$ -DQPSK bit error probability (BEP) of Stüber [STU96]. This formula for BEP is the same expression as for four-phase Gray-coded DPSK for the case of uncorrelated noise samples. The corresponding SEP expressions are also the same. Extending these results to the M -ary case we have an expression for the SEP of π/M -MDPSK that is the same as the SEP of MDPSK for

*This work was sponsored in part by BMDO, managed by the Office of Naval Research under Contract N00014-920-J01761/P00004 and the National Science Foundation under grant number DMI-9423958.

a nonfading Gaussian channel. According to the QA theory, the SEP is of the form

$$P_e(M) = \frac{1}{N} \sum_{k=1}^N P_{e/k}(M | \varphi_k, \beta_k, \beta_{k-1}) \quad (1)$$

where N should be large enough to incorporate errors like ISI patterns due to the filtering at the transmitter and receiver. The conditional SEP is a modified form of MDPSK SEP [PAW82], and is given by

$$P_{e/k}(M | \varphi_k, \beta_k, \beta_{k-1}) = 1 - \int_{-\frac{\pi}{M} + \varphi}^{\frac{\pi}{M} + \varphi} p_\psi(\psi) d\psi \quad (2)$$

where the differential phase error density, $p_\psi(\psi)$, incorporates the instantaneous envelope power deviation factors β_k and β_{k-1} .

The density $p_\psi(\psi)$ is symmetrical so the conditional SEP can be written as

$$P_{e/k}(M | \varphi_k, \beta_k, \beta_{k-1}) = \int_{\frac{\pi}{M} + \varphi}^{\pi} p_\psi(\psi) d\psi + \int_{-\pi}^{-\frac{\pi}{M} + \varphi} p_\psi(\psi) d\psi \quad (3)$$

We can express the conditional SEP with (4) & (5), which are shown at the bottom of this page.

This formula is the most accurate and robust version of the conditional SEP expression, which can be used even for large phase error deviation, even envelope and phase deviations resulting from a Doppler spread Rayleigh fading channel. Since this expression is very calculation intensive due to the need for numerical integration, it results in a long computation time, thus destroying a major advantage of the QA method.

Therefore a faster computation algorithm has to be found. The first step is to simplify the above expression. This can be achieved when we use instead of the envelope power deviations of the present and previous symbol, the average envelope power deviation over the present and previous symbol

$$\beta_k = \frac{\beta_k + \beta_{k-1}}{2}. \quad (6)$$

Another step to simplify the expression above is to find an approximation for (4), since the numerical integration is computationally very intensive, which results in longer computation time.

For a replacement of the expression of the conditional symbol error probability where $|\varphi_k| < \frac{\pi}{M}$, we can use the approximation by Pawula [PAW82]

$$P_E(M) \approx \sqrt{\frac{1 + \cos(\pi/M)}{2 \cos(\pi/M)}} \operatorname{erfc} \left(\sqrt{\frac{E_s}{N_0}} \left[1 - \cos\left(\frac{\pi}{M}\right) \right] \right). \quad (7)$$

This formula can be easily evaluated in a numerical math program like MATLAB®. The fact that the formula is asymptotic means that the error at large E_s/N_0 becomes very small but here is not that large at small E_s/N_0 values. The exact error of the asymptotic formula when compared to the exact analytical expression is characterized in detail in Pawula [PAW84].

Since the approximation is only valid for small phase deviations, this results in a hybrid expression, where the conditional SEP is of the form of

$$P_{e/k}(M | \varphi_k, \beta_k, \beta_{k-1}) = \begin{cases} F\left(-\frac{\pi}{M} + |\varphi_k|, \beta_k, \beta_{k-1}\right) - F\left(\frac{\pi}{M} + |\varphi_k|, \beta_k, \beta_{k-1}\right), & |\varphi_k| < \frac{\pi}{M} \\ 1 + F\left(-\frac{\pi}{M} + |\varphi_k|, \beta_k, \beta_{k-1}\right) - F\left(\frac{\pi}{M} + |\varphi_k|, \beta_k, \beta_{k-1}\right), & |\varphi_k| > \frac{\pi}{M} \end{cases} \quad (4)$$

where

$$F(\psi, \beta_k, \beta_{k-1}) = -\sqrt{\beta_k \beta_{k-1}} \frac{\sin \psi}{4\pi} \int_{-\pi/2}^{\pi/2} \frac{\exp\left(-\frac{E_s}{N_0} \left(\frac{1}{2}(\beta_k + \beta_{k-1})(1 - \sin t) - \sqrt{\beta_k \beta_{k-1}} \cos \psi \cos t\right)\right)}{\left(\frac{1}{2}(\beta_k + \beta_{k-1})(1 - \sin t) - \sqrt{\beta_k \beta_{k-1}} \cos \psi \cos t\right)} dt \quad (5)$$

$$P_{e/k}(M/\varphi_k, \beta_k) \cong$$

$$\sqrt{\frac{1 + \cos\left(\frac{\pi}{M} - \varphi_k\right)}{8 \cos\left(\frac{\pi}{M} - \varphi_k\right)}} \operatorname{erfc}\left(\sqrt{\beta_k \frac{E_s}{N_0} \left[1 - \cos\left(\frac{\pi}{M} - \varphi_k\right)\right]}\right)$$

$$+ \sqrt{\frac{1 + \cos\left(\frac{\pi}{M} + \varphi_k\right)}{8 \cos\left(\frac{\pi}{M} + \varphi_k\right)}} \operatorname{erfc}\left(\sqrt{\beta_k \frac{E_s}{N_0} \left[1 - \cos\left(\frac{\pi}{M} + \varphi_k\right)\right]}\right)$$

$$\text{for } |\varphi_k| < \frac{\pi}{M} \quad (8)$$

and

$$P_{e/k}(M|\varphi_k, \beta_k) =$$

$$1 + F\left(-\frac{\pi}{M} + |\varphi_k|, \beta_k\right) - F\left(\frac{\pi}{M} + |\varphi_k|, \beta_k\right)$$

$$\text{for } |\varphi_k| > \frac{\pi}{M} \quad (9)$$

where

$$F(\psi, \beta_k) =$$

$$-\frac{\sin \psi}{4\pi} \int_{-\pi/2}^{\pi/2} \frac{\exp\left(-\frac{E_s}{N_0} \beta_k (1 - \cos \psi \sin t)\right)}{(1 - \cos \psi \sin t)} dt \quad (10)$$

Because only a few phase errors exceed the value of π/M , the time for computation on average is considerably shorter.

Since we assumed Gray coding of the bits, most k -bit symbol errors contain only a single bit error [PRO95], thus the SEP and BEP are within a good approximation related by $\log_2(M)$. Monte-Carlo simulation was performed to verify that the quasi-analytical method is correct.

Basic System

The system model used is shown in Figure 1. The data is differentially encoded with a modulation index based on the quality of the channel. The symbols are amplitude shaped with a square-root raised-cosine (SRC) shaping function with an excess bandwidth factor of 0.35. The data is converted to an analog signal and passes through an elliptic lowpass reconstruction filter with a cut-off frequency f_c of 35 kHz to fit the IS-130 spectral occupancy requirements. Gaussian white noise is added, and the signal is filtered by the anti-aliasing filter at the receiver, which is also an elliptic lowpass filter with a 35 kHz cut-off frequency. Both lowpass filters model the response of switched

capacitor analog filters. The received signal is converted back to a discrete-time (digital) signal, passes through the matched square-root raised-cosine filter and is differentially detected.

Figure 2 shows the SEP of the basic system for modulation indices of $M = 4, 8, 16, 32$, and 64. In all cases we use four samples per symbol, and eight times oversampling to represent the analog portions of the system. The SEP calculated with the QA method (solid line) and the Monte-Carlo simulation (symbols) is shown. As a reference the analytical values are also shown (dashed line). From the plot one can easily determine that the QA approach is an appropriate method to get simulation results much faster in an AWGN channel. For large modulation indices the degradation increases due to the ISI introduced by the filters and the increased sensitivity to symbol timing errors.

Advanced System

To mitigate the timing error an upsampler after the Square-root raised cosine matched filter at the receiver is used. A value of 8 or 16 upsampling factor produced the best results. In practice this operation would be replaced by an efficient interpolator structure. To combat the error introduced mostly by the reconstruction and antialiasing filters, a zero-forcing equalizer is implemented behind the upsampler at the receiver. The system model of the advanced system is shown in Figure 3. Figure 4 shows the SEP for the advanced system. As one can see the degradation at large modulation indices has decreased considerably. Table 1 shows the improvement of the degradation in E_b/N_0 in dB of the advanced system over the basic system. The degradation in both cases is with respect to ideal system performance in AWGN

Conclusions and further research

High density modulation schemes are indeed more sensitive to implementation imperfections. The QA simulation method is a valid and much faster way to calculate Symbol and Bit Error when these implementation details are being modeled. The zero-forcing equalizer and the upsampler at the receiver are an easy to implement method to mitigate the large performance degradations observed at high modulation indices.

Additional work has been completed in which the QA was method was used to investigate modem performance in a doubly-spread mobile radio channel. These results will be reported in a future paper.

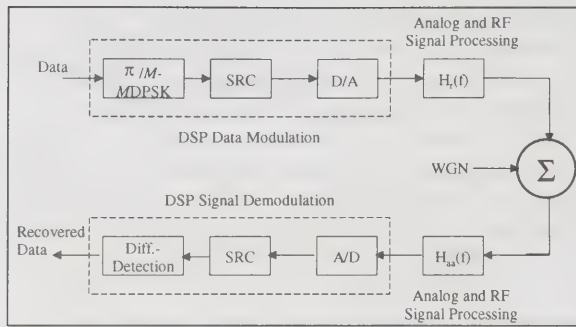


Figure 1: System Model of the basic system

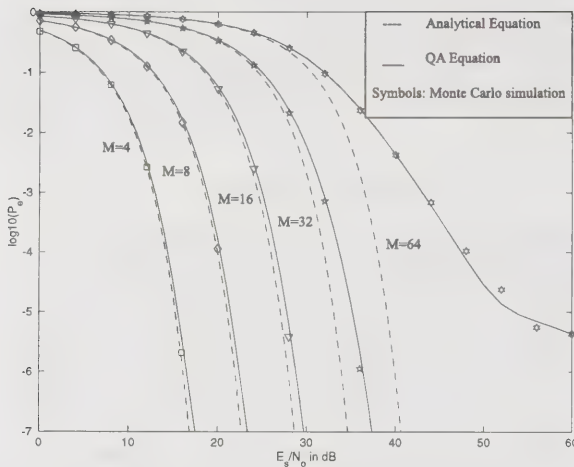


Figure 2: Symbol error probability of the basic system

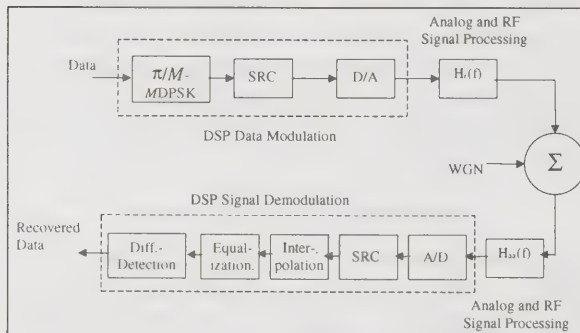


Figure 3: System model of the advanced system

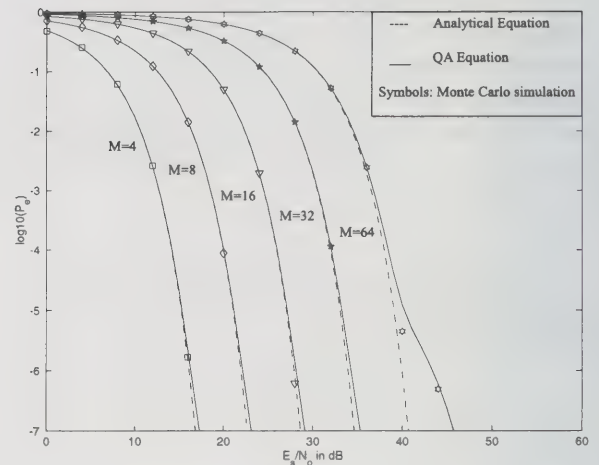


Figure 4: Symbol error probability of the advanced system

Table 1: Improvement of the degradation of E_b/N_0 in dB of the advanced system over the basic system

BEP (P_b)	Improvement of E_b/N_0 degradation in dB over the Basic System				
	$\pi/4$ - DQPS K	$\pi/8$ - 8DPS K	$\pi/16$ - 16DP SK	$\pi/32$ - 32DP SK	$\pi/64$ - 64DP SK
10^{-4}	0.2105	0.2542	0.4755	1.4280	7.2581
10^{-5}	0.1920	0.2929	0.5923	1.8144	10.478
10^{-6}	0.2044	0.3044	0.6407	2.0766	14.919

References

- [JER92] Jeruchim, M., Balaban, P. and Shanmugan, K., *Simulation of Communication Systems*, Plenum Press, New York, Chapters 5 & 6, 1992.
- [PAW82] Pawula, R., Rice, S. and Roberts, J., "Distribution of the Phase Angle Between Two Vectors Perturbed by Gaussian Noise", *IEEE Trans on Comm.*, vol. Com-30, NO.8, pp. 1828-1841, August 1982.
- [PAW84] Pawula, R., "Asymptotics and Error Rate Bound for M-ary DPSK", *IEEE Trans. on Comm.*, VOL. COM-32, NO.1, pp.93-94, Jan. 1984.
- [PRO95] Proakis, J., *Digital Communications*, third edition, Mc Graw-Hill, New York, New York, 1995.
- [STU96] Stüber, G., *Principles of Mobile Communications*, Kluwer Academic Publishing, Boston, MA, 1996.

Transmission Capacity of a Broadband Wireless Radio Link

José Fernandes

Instituto de Telecomunicações - Pólo Aveiro, Universidade de Aveiro

3810 Aveiro - Portugal

Tel.: +351 34 377900 Fax.: +351 34 377901 E-mail: zf@ua.pt

Abstract

The transmission performance of a mobile/wireless broadband digital radio system, using the 60 GHz band, is evaluated for different antenna configurations in two indoor scenarios. It is shown that for the same equalization effort the maximum achievable carrier bit rate strongly depends on the antenna configuration. A CBR of 40-50 Mbit/s in a medium size room and 20 Mbit/s in a large room using biconic horn antennas at MT and at the BST is achievable, while this limit increases up to 100 Mbit/s when using two horn antennas located in the two narrower lateral walls of the room. Moreover, a simple method to estimate the maximum CBR based on the channel time dispersion is presented. Simulation results for several MT positions along the rooms are presented and discussed.

Introduction

The continuous growing need for bandwidth to serve and support broadband mobile/wireless applications will drive the consumer market for broadband products to very competitive prices. Broadband, multimedia and mobility are three key words that will be widely spoken during the next century in the telecommunications field. This tendency is supported by an increasing interest in the Wireless ATM area with very important activities being carried out in Japan and USA and within ETSI and ATM Forum. As broadband application services require large bandwidths, the millimeter-wave band appears to be the best candidate due to the congestion of the lower bands. Thus, two paired bands were provisionally allocated for such type of system: 39.5-40.5 GHz and 42.5-43.5 GHz; 62-63 GHz and 65-66 GHz. The work towards a Mobile Broadband System (MBS) was first initiated by the RACE - MBS project [1] and is being continued by the ACTS SAMBA project (System for Advanced Mobile Broadband Applications) [2].

At high signaling rates, the channel time dispersion can be considered a form of frequency diversity. This is because the signal bandwidth is much larger than the average fade bandwidth, and only a small part of the signal spectrum will be faded at one time. This intrinsic signal diversity limits the signal envelope fading depth at the receiver. As the symbol rate is reduced, the signal bandwidth narrows and, if it becomes comparable with fade bandwidth, a single fade can attenuate a large portion of the signal spectrum, causing a significant envelope fading[3]. This indicates that the time dispersion at high signaling rates provides some form of diversity reducing the envelop fading. However, if the

time dispersion becomes too large an irreducible error rate due to the inter-symbol interference (ISI) is achieved, for a certain equalization effort. The higher the equalizer depth, the higher time dispersion it is able to handle. Since the time dispersion of the radio channel is strongly influenced by the used antenna configuration, it appears that the maximum transmission capacity in an environment strongly depends on the antenna configuration (type of antenna, directivity, location of the base station, etc.) and on the employed equalization effort.

An evaluation of broadband transmission performance, by computer simulation, is presented in this paper considering different antenna configurations in two different indoor environments, using a 4-OQAM modulation scheme and a decision feedback equalizer (DFE) in the receiver. As the characteristics of the radio channel of a pico-cellular based structure of a MBS are strongly influenced by the propagation environment and by the used antennas in the transceivers [4], an appropriate ray tracing tool[5] was used to calculate the channel impulse response (CIR).

Scenarios and Antenna Configuration

One scenario is a rectangular room with dimensions of 24m x 11m x 4.5m and is named Room A. One long side wall is completely windowed and the other walls are mainly made of wood. The concrete floor is covered with carpet and the ceiling is a complex structure made of plastic below the concrete. The other, named Room B, represents a laboratory with dimensions of 60m x 40m x 4m. The two narrower walls are made of wood and concrete. One of the other two walls is complete windowed, the other and the floor are made of wood covering the concrete and the ceiling is plastered.

The definition of an antenna configuration for a given scenario must take into account several parameters, such as the power distribution along the cell, required transmitted power level, channel time dispersion, system transmission capacity, movement freedom of the mobile terminal (MT) and the implementation complexity of the system to mitigate the channel impairments. One configuration consists in the use of a biconic horn antenna at the MT and other at the base station transceiver (BST) centrally placed in the room at a certain height. The other antenna configuration consists in the use of two pyramidal horn antennas located in the middle of the narrower walls of the room tilted downwards relatively to the horizontal direction as shown in Fig. 1. The orientation and beamwidths of each antenna are chosen in order to cover directly 2/3 of the room so as to provide some overlapping area, enabling the MT to choose one/another BST.

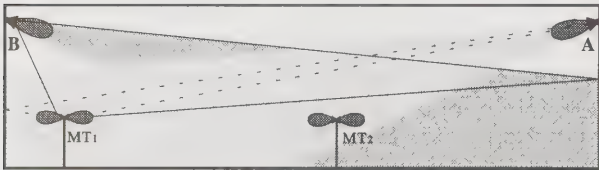


Fig. 1 - Antenna configuration for medium and large size indoor environments (side view).

The used transceiver specifications are identical to those described in [6]. Due to the lack of a channel model able to take into account the propagation environment and the antenna configuration, a set of a thousand CIRs were calculated with a ray tracing tool[5] around each MT spot position in order to simulate the transmission of the same number of information bursts.

Results and discussion

A coherent set of spot positions were defined in both rooms to be representative of the system transmission performance. The typical parameters characterizing the channel impulse response, such as the normalized received power (NRP), rms delay spread (DS) and delay window (DW) in these two scenarios using several antenna configurations can be found in [4]. The bit error rate (BER) was calculated, for the defined set of MT positions as a function of $E[Eb]/No$ for several transmission rates and for the two rooms considering the antenna configurations described above.

Fig. 2 represents BER curves for some MT positions in Room A when using a biconic horn antenna at BST and MT, for a carrier bit rate (CBR) of 40 and 80 Mbit/s, using a DFE with 7 taps in the forward and feedback filters. In this figure, as in the following ones, dotted lines means that MT uses switching diversity and

continuous line means no diversity. The positions represented in figures were chosen in order to give an impression of the transmission performance within room, where the best and worst cases (lower and higher $E[Eb]/No$ values, respectively) were always selected. It can be seen in Fig. 2 how switching diversity reduces the $E[Eb]/No$ requirements for the same BER, this reduction also applies to the transmitting power requirements. The transmission performance for CBR = 80 Mbit/s shows significant differences when compared with 40 Mbit/s, namely for positions 2, 8 and 12. As the symbol time period is 25 ns for 80 Mbit/s, the equalizer depth is ~ 90 ns, then it is able to handle (to equalize) a CIR spanned up to ~ 90 ns. An estimation of the CIR time span can be obtained from DW parameter as explained in [5]. Therefore the equalizer is not able to bring down the BER curve to 10^{-3} (of position 12), which is explained by the fact that the DW is much higher (140 ns) than the equalization depth. Other critical cases were found for positions 2 and 8 which have also high values of DW, respectively 102 ns and 98 ns. These positions are near to the irreducible BER for 80 Mbit/s. This is better illustrated by Fig. 3, where the required $E[Eb]/No$ for several CBRs is shown for some MT positions. It can be observed that as the bit rate increases above 40-50 Mbit/s, the required $E[Eb]/No$ increases. The main error mechanism in this range is ISI, because the symbol time period becomes shorter, decreasing the equalization depth, leading to a higher ISI. Conversely, the required $E[Eb]/No$ increases as CBR decreases below 30-40 Mbit/s. In this range, the main error mechanism is envelope fading because the symbol time gets larger (the signal spectrum bandwidth narrower) and the intrinsic frequency diversity effect observed at high signaling rates becomes less significant. Another fact that confirms this idea, is the higher improvement obtained with antenna diversity for lower bit rates.

Fig. 4 represents the transmission performance for three MT positions and demonstrates that the coverage is not so uniform as it is with a pair of biconic horn antennas. As these three positions represent the best and worst cases in the coverage area of the BST antenna, it can be concluded that a minimum CBR of 100 Mbit/s is guaranteed, and in some positions it can reach 500 Mbit/s. This illustrates the influence of the antenna configuration on propagation channel parameters (e.g., the value of DW ranges from 13 to 70 ns) and on the transmission performance (CBR ranges from 100 to 500 Mbit/s).

The performance improvement of this antenna configuration becomes even more evident if we compare the required transmitted power. Considering a receiver with a noise figure of 6 dB, a system margin of

4 dB and a CBR of 40 Mbit/s, the required transmitted power is 10 dBm for the configuration that uses two biconic horns, while it reduces to 3 dBm with the laterally placed pyramidal horn antennas.

A similar study was performed in the Room B. The Fig. 5 represents the $E[Eb]/No$ as a function of CBR required to have a BER of 10^{-3} when using a biconic horn antenna at BST and MT. It is clear from the figure that the positions 18 and 19 impose the minimum CBR. This complies with the fact that they are near the room center and have the highest time dispersion, DW is about 270 ns, which gives a bit rate near 25 Mbit/s ($T_s \cong 77$ ns). On the other hand, the position 1 has the lowest value of DW parameter, 100 ns, which gives a CBR of 70 Mbit/s ($T_s \cong 28$ ns), as shown in Fig. 5.

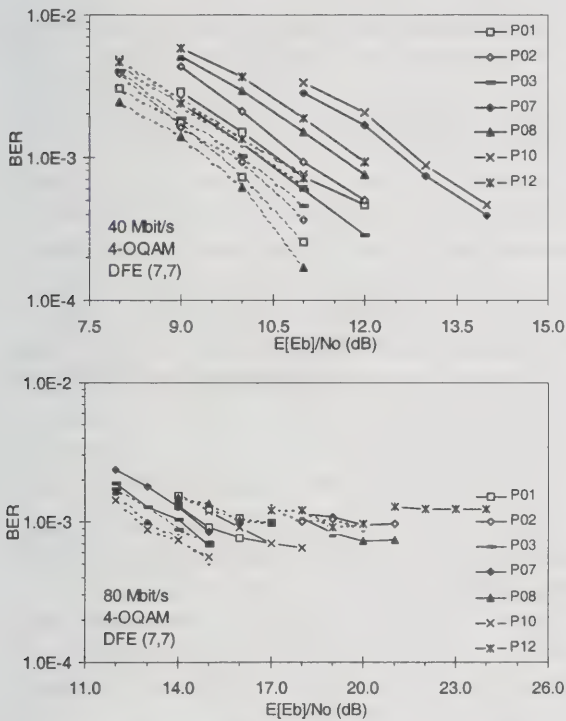


Fig. 2- BER curves for some MT positions: 40 and 80 Mbit/s

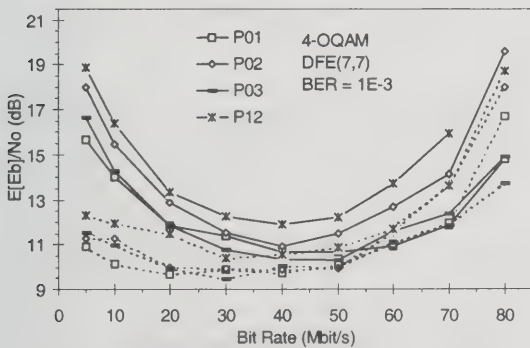


Fig. 3 - Required $E[Eb]/No$ as a function of CBR for a centrally placed BS and a biconical horn at MT ($DW_{P01}=90$ ns, $DW_{P02}=102$ ns, $DW_{P03}=100$ ns, $DW_{P12}=140$ ns).

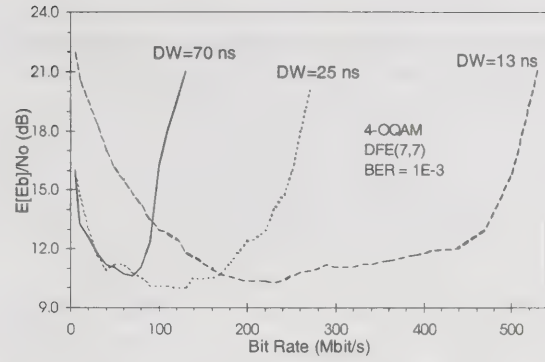


Fig. 4 - Required $E[Eb]/No$ as a function of CBR for a laterally placed BS and a biconical horn at MT in three different positions.

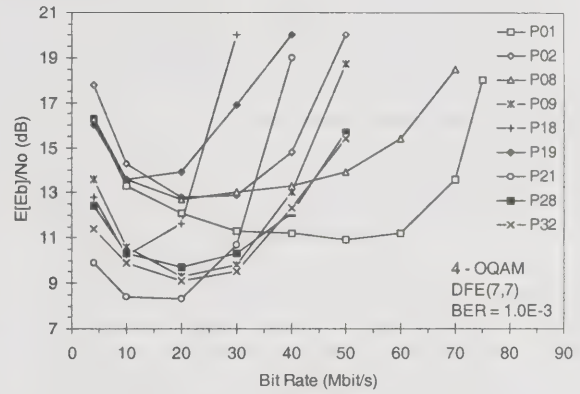


Fig. 5 - Required $E[Eb]/No$ as a function of CBR

The Fig. 6 represents the $E[Eb]/No$ as a function of CBR required to have a BER of 10^{-3} when the BST located in the point B is used. The worst performance was obtained for positions 1, 2 and 9, but these positions are not supposed to be served by the BST located in B, but in A. The transmission performance exhibited in the other positions in the figure, indicates that a minimum of 100 Mbit/s can be achieved in the coverage area of the BST, which is similar to that achieved in Room A. However, when the biconic horns are used at both stations, the achievable CBR is about 20 Mbit/s, nearly half of the achieved in Room A. This illustrates the impact of the room size, which increases the channel time dispersion, and that the configuration that uses two BSTs in the lateral walls is adequate for medium and large size rooms.

A detailed study was also performed in the two rooms using a six sector antenna at both stations. As each antenna has just one sector active at one time, a significant reduction of the channel time dispersion is achieved which means that CBRs higher than 100 Mbit/s are possible. Several signal processing techniques can be employed in order to increase the upper limit imposed on bit rate due to multipath propagation, but they can not cope, either due to cost

and/or complexity, with very large channel time dispersion. It appears that the best solution is to use adequate antenna configurations in order to reduce the time dispersion and the required transmitted power.

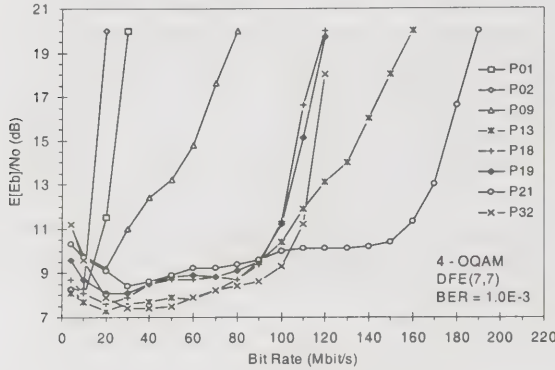


Fig. 6 - Required $E[Eb]/No$ as a function of CBR

Estimation of the Equalization Depth

The channel time dispersion causes ISI and in the limit imposes an irreducible BER. One important issue is to relate the channel time dispersion with the achievable bit rate per carrier, but for that we need to estimate the equalization depth based on equalizer configuration. The Fig. 7 depicts part of the several simulations performed using a DFE(N, M), for different values of N and M , respectively the number of taps in the forward and in the feedback filter. This figure shows that, for the same $E[Eb]/No$, the bit rate increases with the number of taps of the DFE. Thus, the required transmitted power level is the same, while the complexity of the equalizer is increased.

For the same $N + M$ the performance characteristics are different: with a DFE(N, M) where $M > N$ and other DFE(N', M') with $N' > M'$ and $N + M = N' + M'$, the first configuration performs better. Since $M \geq N$ the equalization depth (Ed) can be approximated by

$$Ed = \frac{N + M}{2} TSp \quad (1)$$

with TSp being the tap spacing of the equalizer filters, which is, generally, a fraction of the symbol time. In case of $N \geq M$ the excess number of taps has nearly half of the performance as compared in the feedback filter, then

$$Ed = \left(\frac{3}{4} M + \frac{1}{4} N \right) TSp \quad (2)$$

Once the equalization depth known, the CBR can be estimated through the use of DW parameter, which must be lower than the equalization depth.

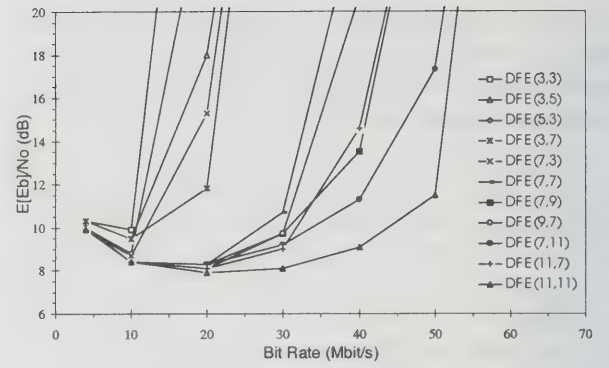


Fig. 7 $E[Eb]/No$ as a function of carrier bit rate for different values of N and M .

Conclusions

The broadband transmission performance of indoor radio mobile/wireless transmission was evaluated. The results have shown that it is possible to achieve a CBR of 40-50 Mbit/s in a medium size room and 20 Mbit/s in a large room using biconic horn antennas at MT and at the BST, while this limit increases up to 100 Mbit/s when using two pyramidal horn antennas located in the two narrower lateral walls of the room. This suggests that an antenna configuration should be selected according to the power distribution along the cell and to the channel time dispersion. This work highlights how important is to select an adequate antenna configuration in an environment.

References

- [1] L. Fernandes, "Developing a System Concept and Technologies for Mobile Broadband Comm.", IEEE Personal Comm. Magazine, Feb. 1995.
- [2] M. Dinis, V. Lagarto, M. Prögler and J. Zubrzycki, "SAMBA: a Step to Bring MBS to the People", ACTS People Mobile Comm. Summit '97, Aalborg, Denmark, Oct. 1997.
- [3] M. Gibbard, G. Morrison, A. Seasy and M. Fat-touche, "Broadband TDMA over the Indoor Radio Channel", Electronics Letters, Vol. 30 No 4, Feb. 1994.
- [4] J. Fernandes, O. Afonso and J. Neves, "Impact of Antenna Set-up and Arrays on Mobile Radio Systems", ICUPC'95, Japan, Nov. 1995.
- [5] J. Fernandes, J. Neves and P. Smulders, "Mm-Wave Indoor Radio Channel Modelling vs. Measurements", Wireless Personal Comm. Journal, vol. 1, 1995.
- [6] A. Gusmão, J. Fernandes and Rui Dinis, "Performance of MBS Radio Transmission", RACE Mobile Workshop, Amsterdam, May 1994.

Cellular System Performance Prediction

Ron Rudokas, Member, IEEE

Abstract – Shannon's theorem combined with cellular system relationships between deployment, operational parameters and S/I form a relationship between system capacity and system re-use. Several enhancements are derived that more faithfully represent typical operational behavior of cellular systems. Results indicate that there is an optimum re-use factor ($k=4$) that maximizes system voice capacity.

Introduction

Cellular network performance estimates, Lee [1], are enhanced to produce a simple relationship that includes the effects of antenna pattern, location error, tower height, power control, system loading and voice activity. In combination with Shannon's theorem a robust optimum network configuration for maximum system capacity is found.

Shannon's Theorem

The Shannon-Hartley theorem [4] defines a limit to the data carrying capacity of a noisy channel.

$$C \leq W \log_2 \left(1 + \frac{S}{N+I} \right) \quad 1$$

C Channel capacity (bps/Hz) r Interference
W Channel bandwidth (Hz/ user) N Noise

This ideal can never be achieved in a practical system. Figure 1 shows deployed cellular system performance with Shannon's curve. The best fit is consistent at 19 dB less than the Shannon limit. This operational loss, Δ_R , consists of margins for fading, coding and equipment realization.

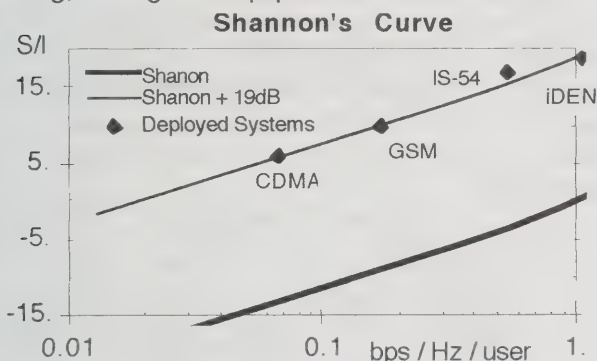


Figure 1. Shannon's curve with live system data.

Table 1. Deployed technology parameters.

Typical Existing Technology Parameters						
System	Format	BW kHz	k	Field S/I	kbps per ch	User / ch
IS-54	QPSK	30	7	18	48.6	3
iDEN	16-QAM	20	7	19	64	3
GSM	GMSK	200	4	10	270.8	8
CDMA	SS	1228	1	6	1228	15

Deployed technology data is based on channel data rate per user and channel bandwidth. In CDMA

the number of users per channel vary based on conditions and the quality of service. A nominal performance level is chosen that reflects typical published performance. Comparison to deployed technologies is critical in relating theory to practical implementation.

The critical cellular systems capacity metric is voice capacity per cell site, Erlangs per site, embedded in a system of sites. Equation 1 is rewritten in terms of total system bandwidth and the re-use factor, k by substituting equation 2 for W . Channels used for system control are removed before calculating the Erlang capacity of the site.

$$W = \frac{BW_{sys}}{k'(Q_v + Q_c)} \quad 2$$

$$Q_v \leq \left[\frac{BW_{sys}}{k'} \frac{\log_2 \left(1 + \frac{S}{N+I} \right)}{C_u} \right] - Q_c \quad 3$$

$$\frac{Erl}{site} \approx Erl(B, Q_v) \quad 4$$

omni $k' = k$ sector $k' = 3k$
 Q_v voice channels Q_c control channels
 BW_{sys} system bandwidth C_u bit rate per user
 $Erl/site$ Erlangs per site at B % blocking

Equation 3 shows that capacity optimization occurs on several competing levels. Maximizing system S/I or bandwidth to achieve the greatest capacity per channel or minimizing re-use or data rate to achieve the maximum channels per site.

Cellular System S/I Estimates

For $k=1$ there are two possible configurations. Each cell may use the same three frequencies with a unique frequency in each sector. This is the logical extension of the sectorized $k=7$ pattern. Interferors are nominally one cell site diameter away from the user. In a second configuration a single frequency is used in each sector of each cell. In this case interferors are nominally at one cell site radius and the configuration is nearly identical for both omni and sector systems.

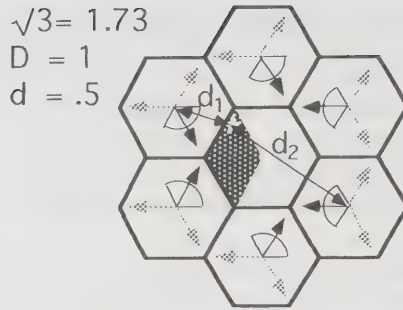


Figure 3. $k = 1$, re-use.

The use of ideal S/I analysis [1] is problematic for small k due to the discrepancies that occur between D/R , typically used for S/I estimates, and d/R the distance between the user and the interferor. S/I is determined by estimating the power of the desired signal relative to interference. A sector design, for $k=3$ and 4, has three first tier interferors in the antenna main lobe. Equation 5 illustrates the relationship between S/I and the relative separation between sites and the propagation constant, γ .

$$\frac{S}{N+I} = \frac{S}{N + \sum_{j=1}^m I_j} \approx \frac{R^{-\gamma}}{N + \sum_{j=1}^m d_j^{-\gamma_j}} \approx \frac{\text{desired signal}}{\text{noise + interference}} \quad 5$$

d_j distance to the interferor R cell radius

This cumbersome relationship is often simplified by assuming that γ (usually 40dB/decade) is the same for each interferor and that d , the distance from user to interferor, is equal to D , the distance from cell site center to center. In sector systems an ideal antenna is assumed where only two sites in the main antenna lobe, are considered. While a good estimate for re-use patterns of seven and greater the large error between d and D , 60% at $k=1$ figure 3, result in optimistic S/I estimates for small k .

Equation 6 captures the details of operational cellular systems by utilizing a table lookup for each d , perturbing each d to include location error, calculating a γ for each site based on height and including interferors scaled by the antenna gain.

$$\frac{S}{N+I} \approx \frac{G R^{-\gamma}}{N + \sum_{j=1}^m G_j d_j^{-\gamma_j}} \quad 6$$

G_j antenna gain in the direction of the j th Interferor
 d_j interference distance perturbed by location error
 $m=18$ both 1st and 2nd tier interfering cell sites.

Antenna Height

Antenna height and local topography dramatically affect propagation characteristics. Deployed systems use a variety of tower heights due to

terrain, available sites and coverage concerns. A histogram of antenna height, for a deployed system, figure 4, shows a mean height of 195 feet with a standard deviation of 90 feet.

Okumura and Hata computed an approximate relationship between tower height and propagation constant. γ is given by [2]:

$$\gamma_j \approx 44.9 - 6.55 \log(h_b)$$

h_b tower height above average terrain (HAAT).

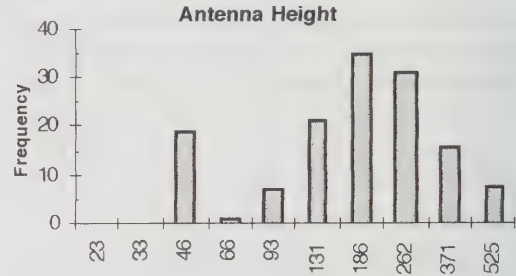


Figure 4. Antenna height from field data.

To achieve a -40dB/decade propagation constant, often used in capacity analysis, a tower height of 18 feet is required. From figure 4, a more pragmatic tower height is 195 feet resulting in a γ of 34 dB / decade. Surrounding structures, at or above the antenna, cause γ to increase rapidly.

Antenna Pattern

The use of ideal anechoic chamber antenna patterns results in optimistic S/I estimates. When the degradation of mounting structures, ground clutter, other antennae or nearby reflectors is considered, directional antennae adopt a somewhat offset omni shape, figure 6. Individual sites can vary dramatically. Field data show that sector sites have significant back and side lobe degradation.

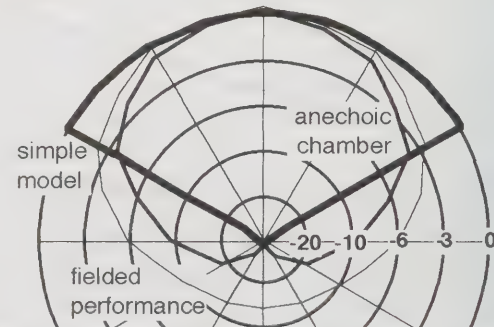


Figure 6. Typical antenna patterns used in analysis.

Location Error

S/I estimates are further refined by allowing cell site location error. Location error is estimated by analyzing the distribution of deployed cell site radii, figure 7. The radius is defined as one half the average distance to the closest three sites. Mean

radius is 3.36 miles with 95% of the sites distributed at less than $\pm 12\%$ about the mean.

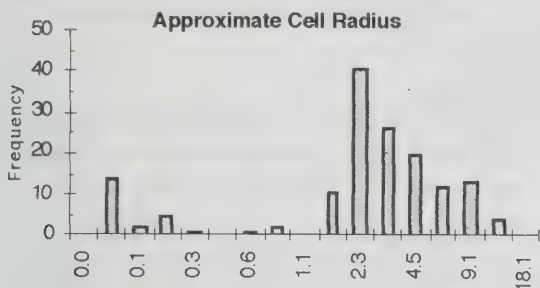


Figure 7. Deployed cell site radius (miles).

Cell site location errors have a greater affect on net S/I, in tight re-use patterns, since the errors are much larger relative to the re-use distance. Systems that utilize a small k factor (or a lower S/I) are more sensitive to design and optimization.

Telephony Traffic

In cellular systems the net system downlink and uplink S/I is a function of the loading and activity in the system. Under most conditions system S/I is significantly greater than predicted since the transmitter is intermittently active. An estimate of the S/I improvement is made through the use of the Erlang equation and system blocking requirements.

As the number of servers decreases a greater percentage of those servers must be held in reserve to insure that a certain blocking rate is achieved. As the number of servers grows the statistical likelihood of one call ending just as another is placed increases and a smaller percentage of the servers are idle. The utilization of each channel and system S/I improvement is [3]:

$$\Delta_U = \frac{1}{U(n)} = \frac{n}{Erl(B,n) (1 - B)}$$

U(n) percentage of time that each channel is active
Erl(B,n) offered traffic B % call blocking
 Δ_U S/I improvement n trunked servers

As the number of channels per cell site and the usage per channel increase, S/I decreases.

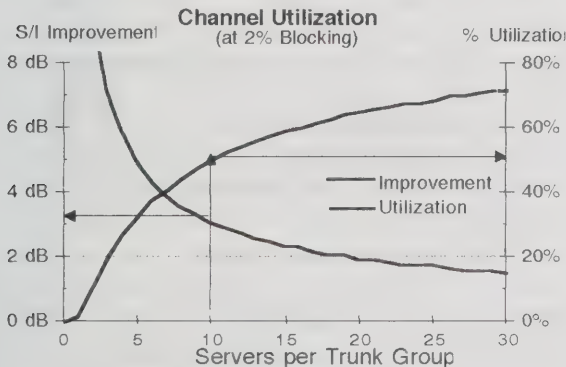


Figure 8. Channel utilization

Most TDMA type technologies, (IS-54, GSM, iDEN) are always active since there may be multiple users per RF channel. In this case there is no improvement to downlink S/I. Uplink S/I improves by the amount predicted in figure 8 since the subscriber unit transmits only when in use.

In a similar fashion S/I is improved if voice activity is used to control the “on” time of the transmitter. During periods of silence no transmissions occur reducing the average system interference. At 50% voice activity [6], system S/I improves 3dB.

$$\Delta_{VA} \approx \frac{1}{A_v} \tag{9}$$

Δ_{VA} % time of active voice

Transmit Power Control

System S/I suffers when excessive power is transmitted. A common feature of many cellular technologies is the reduction of transmit power for users close to the cell site. The improvement is evaluated by considering the difference between the transmit power with and without power control in a donut around the cell site. Simple calculus results in equation 10. At a γ of 35 dB / decade S/I improves 3dB.

$$\Delta_{PC} = \frac{\pi \gamma}{\gamma + 2} \tag{10}$$

Re-use Summary

Combining equations 6, 7,8,9 and 10 provides an estimate of median net operating S/I.

$$\frac{S}{I_{Sys}} \approx \left(\frac{GR^{-\gamma_h}}{N + \sum_{j=1}^m G_j d_j^{-\gamma_{h_j}}} \right) \Delta_{PC} \Delta_{VA} \Delta_U \Delta_R \tag{11}$$

In the effort to deploy digital cellular systems many organizations have revisited the AMPS 18dB S/I criteria. This testing has shown that an AMPS call, at a 18dB faded S/I, results in a typical MOS score of less than 2, which is judged unacceptable by most customers. A faded S/I of 21dB is required to achieve a MOS score near 4, which is acceptable.

Early experiences set expectations for the quality of cellular networks. Initial cellular systems were deployed as omni systems for mobile coverage with few channels per site. Under these conditions the system S/I greatly exceeded the design S/I due to the low utilization of channels, dispersed deployment and relatively uniform tower height (all high sites). Customers grew to expect the quality of service that this type of system offered.

As the customer base grew, more channels and sites were deployed, reducing the S/I improvement due to system traffic and voice activity, as predicted in equations 9 and 10. In most systems the original high sites were retained hampering efforts to improve system S/I [5]. This brought busy hour S/I to the 18dB level ($MOS \leq 2$) with the expected customer dissatisfaction in the quality of service.

Results

Combining equations 4 and 10 produce the following figures using these parameters:

System bandwidth 12.5MHz bps/user 16000

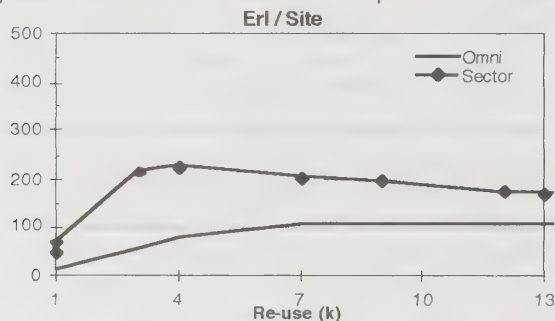


Figure 9 Total system capacity vs. k

These surprising results are entirely reasonable when the factors that apply to the capacity of a cellular "system" are considered. At small k the dominant factor is the interference generated by nearby interferors. For re-use less than 4 the bandwidth per user increases dramatically in order to compensate for the increased interference. The increasing bandwidth reduces the number of available channels and the Erlang efficiency, causing a dramatic drop in system capacity. The two sector data points, at $k=1$, represent the one RF channel per site and three RF channel per site cases; one channel per site having slightly less capacity due to increased interference. For any reasonable combination of parameters the peak system capacity occurs at a re-use of 4 with a modulation efficiency of 1.3 bps/Hz. The system S/I is 21dB due to the effects of ΔV_A , ΔP_C , and ΔU .

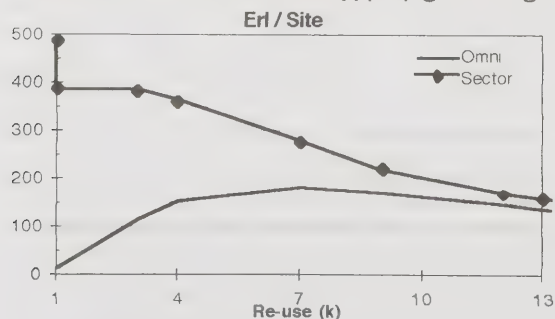


Figure 10. Total system capacity vs. k (ideal antenna pattern, 40dB/decade propagation)

Figure 10 is the more common result that occurs with ideal antenna patterns and a 40dB/decade γ . The system configuration and d/R ratio's are the same as in figure 13.

Conclusions

A relationship comprising Shannon's law, cellular S/I and deployed system performance shows that over a broad range of conditions a modulation format optimized for a re-use of 4 delivers the greatest possible cellular "system" capacity. Large changes in parameters do not change the location of the maximum. At $k=4$ the reduction of capacity, due to fewer voice channels per hertz and poor Erlang efficiency is balanced by an increase in system capacity due to the greater re-use of RF channels. This capacity peak is optimized relative to the entire system rather than one site or one modulation parameter.

Although these relationships do not describe any existing technology they do provide an easy method of estimating the effectiveness of various technology proposals. For example, since the dominant contributor to the reduction of capacity, for $k \leq 3$, is degraded antennae patterns an opportunity exists for transmit and receive smart antenna technologies to alleviate this problem.

References

- [1] C. Y. Lee, Mobile Cellular Telecommunications McGraw Hill, 1995.
- [2] Raymond Steele, Mobile Radio Communications, IEEE Press, 1992.
- [3] Aaron Kershenbaum, Telecommunications Network Design Algorithms, McGraw Hill, 1993.
- [4] Bernard Sklar, Digital Communications, Fundamentals and Applications, Prentice Hall, 1988
- [5] Terry Benz, Cellular system designer and operator, private communication.
- [6] Ryuji Kohno, et al, Spread Spectrum Access Methods, IEEE Communications Magazine, Jan. 95 Vol.33 No 1.

Bibliography

Ron Rudokas received his MS EE from the University of Illinois at Champaign in 1976. He has held senior design, development and operations positions at TRW, PacTel and NexTel. He holds a patent for RF based pre-call fraud control. Joining OneComm in 1994, led the operational development and deployment of the iDEN system. Ron is a consultant specializing in wireless technologies and systems.

RRudokas@ix.netcom.com.

Toward the 60 gm Wireless Phone

David H. Smithgall
Lucent Technologies
Box 900, Princeton, New Jersey 08540
dsmithgall@lucent.com

Abstract

This paper discusses the complementary relationships among component technology, physical design, energy consumption and battery design in reducing the size of wireless terminals.

Introduction

The wireless handset terminal market continues to be driven to reduce size, weight, and cost, while continually increasing functionality. The minimum size for phones in the market today depends upon network technology, but those designed for the North American TDMA or CDMA digital standards typically weigh in the range 175 - 225 gm. For these phones functionality may be measured in terms of features such as memory dialing, message services, paging, or voice dialing. It also includes the duration of talk and standby times which are expected to increase with each succeeding generation of phone models. The designer is faced with the challenge of reducing size and cost while increasing the energy efficiency of the terminal. Figure 1 illustrates the weight trend for analog terminals, and the weights of recently introduced digital phones. Projecting the weights of the digital phones at more aggressive rates indicates that a 2/3 reduction might be achieved in 2-5 years. But what are the implications for batteries and circuit energy efficiency?

Considering the phone in terms of energy use, energy supplied by the battery is either transmitted according to a specified standard, or consumed by the circuit

components within the phone. Acceptable battery technologies are practically limited to NiMH or Li Ion today, which have limited energy/weight and energy/volume ratios. Li Ion technology produces

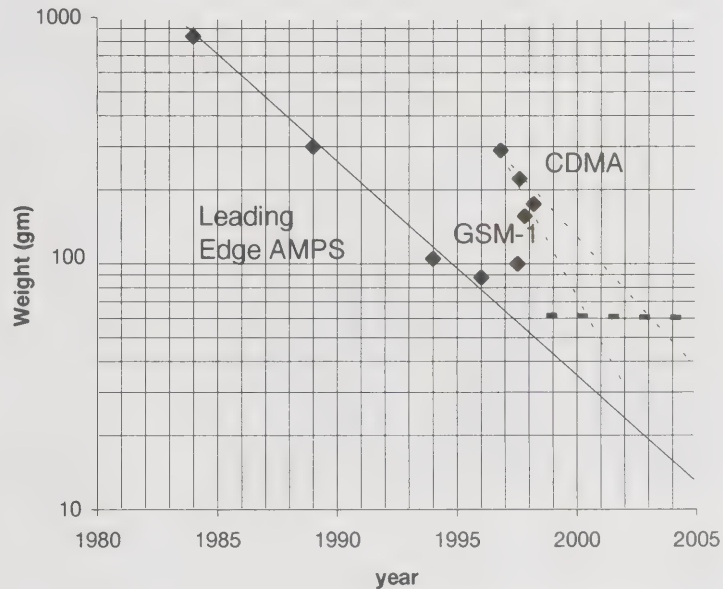


Figure 1 - Trend in Wireless Terminal Weight

approximately 50% more energy/weight than does NiMH technology, and is the preferred choice for lightweight implementations, though at a 2:1 cost penalty.

Since the available energy is defined by battery technology and the fraction of that energy radiated from the antenna is defined by standards, energy efficient design involves minimizing power consumption of circuits within the wireless terminal itself.

Wireless Terminal Model

The wireless terminal can be modeled as the five blocks shown in Figure 2. The internal circuits in the phone are divided into the radio (RF) and baseband circuits. The user interface provides for interaction between the user and the phone. In its most elementary form the user interface consists of a keypad and microphone for input and a speaker and display for output. In addition, there will may be function keys, an accessories/charging connector, and an alerter. From the standpoint of reducing cost and size, the functionality of the user interface typically provides a constraint based upon intended market, features and user friendliness. The display may be a simple one-line character based display for the number entered, or a bit mapped graphic display for multi-language messaging. The energy source is typically a rechargeable battery, the size and capacity of which are determined by the required talk and standby time and the circuit design. The mechanical package, which supports the battery and the internal circuits, can be reduced to the extent allowed by those components and is further constrained by the user interface requirements.

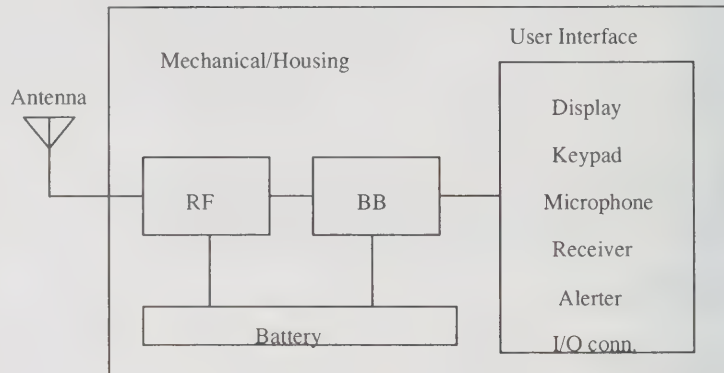


Figure 2 - Wireless Terminal Model

While the radio circuit can benefit from functional integration, progress toward integrating high frequency components is much slower than baseband technology. The radio is primarily an analog circuit, where nonlinear processes and components require careful board layout and filtering. In this circuit, size and cost reduction may benefit from architecture changes as much as from integration. For example, replacing the present 2-3 stage demodulators with direct demodulation to baseband would significantly reduce cost and size of this circuit. Some benefits could also be realized by replacing the large ceramic and SAW filters with novel components such as thin film resonators [1].

RF and Baseband Models

The RF and baseband circuits follow different cost and size reduction models due to the technology available. With the exception of the audio circuit, the baseband is essentially a silicon digital technology which can follow the miniaturization trends of the semiconductor industry. The criteria for implementing a functional integration or component size and energy reduction are the stability of the design and the capabilities of semiconductor suppliers. In addition, while mixing semiconductor technologies for memory, processors, and mixed signal

Energy Model

The battery capacity (W-hr) required to achieve a given talk or standby performance (hr), is determined from the average power (W) consumed by the phone. It is convenient to model the phone as three elements. The first comprises the baseband circuits which were also identified as a physical model. The second element comprises the RF integrated circuits and associated components exclusive of the power amplifier. The third element, referred to as the "transmit path" includes

the circuit elements between the power amplifier and the antenna (Figure 3). The passive elements provide a fixed attenuation, and the efficiency of the power amplifier is the design parameter. In an aggressive design for a CDMA wireless terminal with an average power consumption of 1 W in the talk mode, as illustrated in Figure 3, energy consumption is allocated among the three elements.

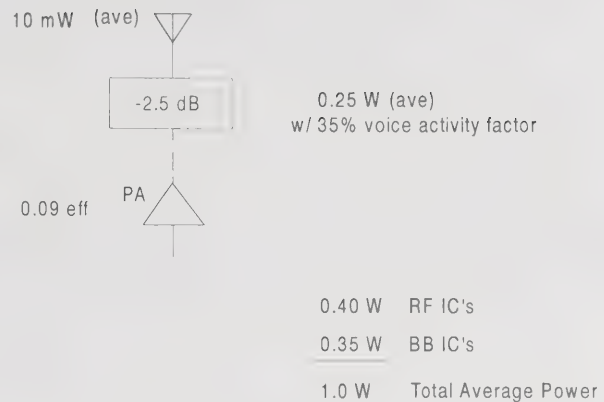


Figure 3 - Transmit Path Energy Model

Figure 4 illustrates the effect of migrating to lower voltage circuit components to achieve a design goal which requires a reduction in average energy consumption to 0.45 W. The greatest impact is seen in the primarily digital baseband circuit. The benefits are reduced in the RF circuits with the dominance of analog signals, and there is essentially no benefit in the transmit path. Figure 4 indicates that the desired 55% reduction in energy consumption could be realized with advances in integrated circuit technology.

Summary

Wireless terminal size, weight, and cost will continue to be reduced through a combination of evolving low-voltage, high density semiconductor technologies and architectural changes. The size and weight of the phone will ultimately be determined by the user interface requirements, for the feature functionality, human factors, and the useful talk and standby times achieved.

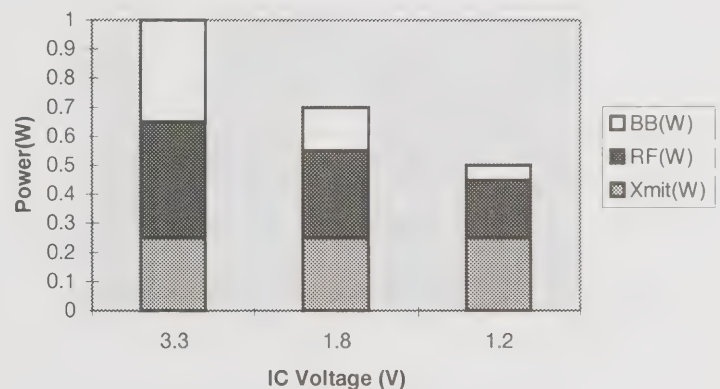


Figure 4 - Power Dependence on IC Voltage

Reference

1. K. M. Lakin, "Recent Advances in Miniature Filters", Proceedings of the Wireless Communications Conference, Boulder, CO, August 19-21, 1996, pp. 33-34.

Comparison of Antenna Diversity Schemes

B. Gavilanes-Loureiro¹, I. De Coster¹, E. Van Lil^{1,2}, F. Pérez-Fontán³

¹K.U.Leuven, div. ESAT/TELEMIC, Kard. Mercierlaan, 94; B-3001 Heverlee; Belgium

☎+32/16/321113; Fax: +32/16/321986; ✉emmanuel.vanlil@esat.kuleuven.ac.be

²Fund for Scientific Research (Flanders)

³Univ. of Vigo, Spain.

Abstract

This paper compares 4 kinds of simple diversity schemes for an indoor scenario at 2.4 GHz: polarisation, dual antenna, antenna and polarisation, and triple antenna for an indoor scenario, all combined according to 5 different techniques. Useful parameters are extracted statistically. We conclude that the power combining method gives the best results, and that the extra complexity of the 3 antenna case is not always justified from a cost viewpoint.

Key words: Antenna diversity, polarisation diversity, indoor propagation.

Introduction

Even if the foundations for diversity studies have been laid out long ago [1], it took until both fast computers and efficient computation schemes [2] were available before simulations of complex scenarios was possible. In this paper simulations are given for 4 different diversity cases.

a. The first uses polarisation diversity and a dual polarised antenna. Its advantage is that the antenna does not take up more space than a single antenna.

b. The second combines the signals of 2 antennas, where the distance between the two antennas was chosen such that the correlation was smaller than 0.5 (in this case one wavelength).

c. The third combines antenna and polarisation diversity.

d. The fourth uses 3 antennas, not placed like in [1] on an equilateral triangle, but forming a simple equidistant linear array with an inter-element spacing of one wavelength.

A few representative but still simple test cases have been computed. The first was a reference room taken from the literature [3]. Because the level of the different components of the fields were significantly different due to the perfectly aligned vertical walls of the objects, and this is not encountered in practical measurements [4], we can better illustrate the advantages of diversity schemes by using a more futuristic looking chamber (Fig. 1). The position of the transmitter is indicated Fig. 1, and the receiver moves along a line where the field is sampled with a high resolution (961 points = 1 point per $\lambda/20$) partly in LOS conditions and partly in NLOS conditions.

Field computations were made with a GO procedure and not with the improved model that we have developed [2], because this will not change significantly the statistics and hence our conclusions.

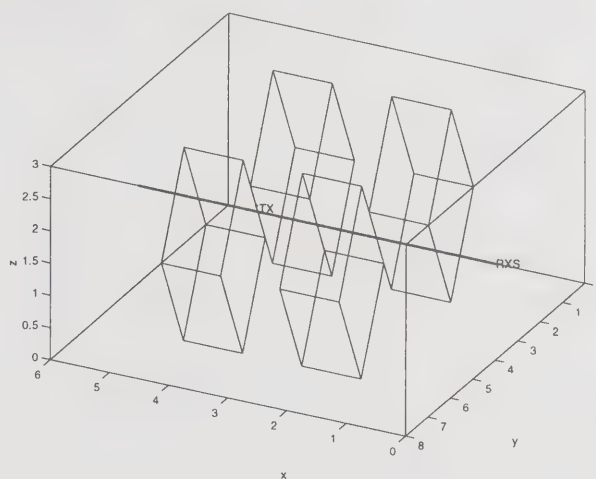


Fig. 1: Chamber with oblique objects.

We can also see from the fields (Fig. 2) that the central part is in LOS condition, and that the other polarisations are relatively small there, while the other parts are in NLOS conditions and do have significant cross-polar levels.

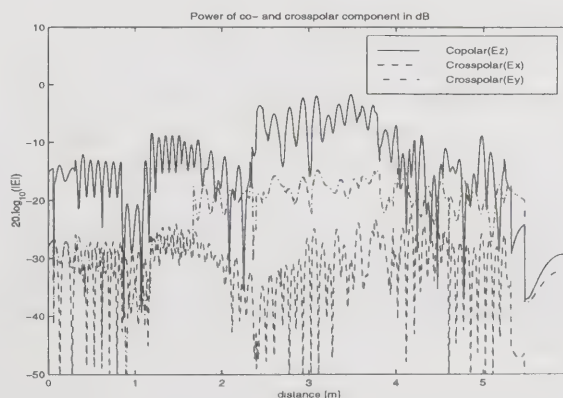


Fig. 2: The 3 field components along the trajectory.

Results and interpretation

5 combining methods were considered:

1. The first combines all elements in phase in a power combiner. This gives either a broadside array for the antenna combining methods or a linearly polarised antenna for the polarisation combining methods.
2. The second combines all elements with a 90° phase shift in a power combiner. This tilts the beam in case of an array, or makes a circularly polarised antenna in the polarisation combining case.
3. The third chooses the largest signal.

4. The fourth adds the decoded signals together (MRC). This corresponds with a factor of \sqrt{M} with respect to a mere switching between the antennas ($M=2$ in cases a, b and c; $M=3$ for case d).

5. The fifth combines the powers of the signals.

Some significant statistical parameters are summarised in Table 1. Not only the increase of the mean levels, but specially the minimum levels show that the power combining method is the best. The next rows show the improvement in signal level obtained for respectively 99%, 90% and 80% of the cases. (or that is not obtained in 1%, 10% and 20% of the cases).

Table 1: Comparison of the signal levels of the polarisation combination methods (x- & z-components).

	Vertical	Horizontal	Linear	Circular	Maximum	Amplitude	Power
Mean	-12.2	-21.4	-15.5	-13.9	-12.0	-12.6	-11.5
σ	-14.5	-26.2	-18.5	-16.1	-14.7	-16.2	-14.7
Min.	-40.5	-62.0	-62.6	-43.7	-37.1	-34.4	-34.4
Max.	-1.6	-14.4	-5.0	-3.4	-1.6	-3.1	-1.5
99%			-22.0	-0.3	0.4	3.0	3.0
90%			-2.4	-1.3	0.9	1.7	1.8
80%			-2.2	-0.8	2.4	2.5	3.4

A graphical representation of the CDF's is seen for the case of polarisation diversity in Fig. 3. The 2 RF combining methods even give worse results than the single-antenna case. The vertical scale used is the same as in the representation of the Nakagami distributions in [5], and has the advantage that a Rayleigh CDF is represented as a straight line. As a check, both a Rayleigh and a Rice CDF fitted to the signal of the reference antenna (Copolar) is also represented on the same plot. For fitting the following formulas were used:

$$E\{A^4\} = 8\sigma^4 + 8\sigma^2 K^2 + K^4$$

$$E\{A^2\} = 2\sigma^2 + K^2$$

from which: $2\sigma^2 = E\{A^2\} \pm \sqrt{2E\{A^2\} - E\{A^4\}}$

and: $2K^2 = E\{A^2\} - 2\sigma^2$.

In this case the determinant is negative, leading to a minimum square error at $2\sigma^2 = E\{A^2\}$, and $K=0$. This explains the 2 parallel lines (Rayleigh distributions, but with a different sigma).

The match to the curve of antenna 1 is not perfect because the signal has both LOS (Rice) and NLOS (Rayleigh) parts. The number of points is sufficient to allow comparisons on levels above 99.9%.

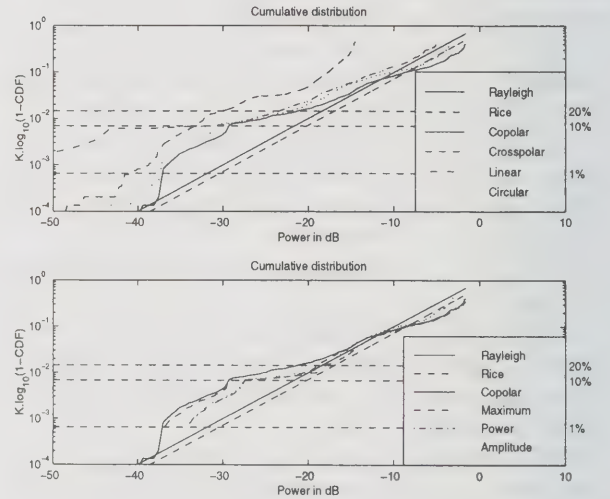


Fig. 3: Cumulative distribution functions of the polarisation diversity.

From the amplitude correlation plots (Fig. 4) one can see that the minimal separation for the 2-antenna case (where the correlation is less than 0.5) is a little bit more than 1 wavelength. This is the value that we have taken to apply the 5 diversity schemes. The corresponding CDF's are given in Fig. 5. The same conclusions do apply: the power combination gives the best result, but the improvement is better than in the polarisation diversity case.

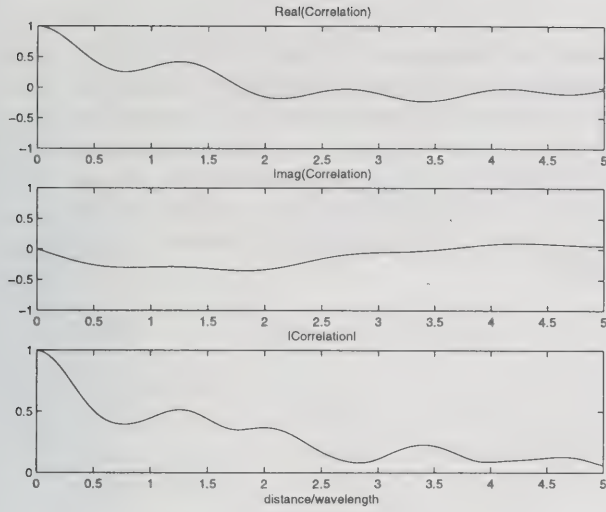


Fig. 4: Correlation plots for the 2-antenna case in function of the distance.

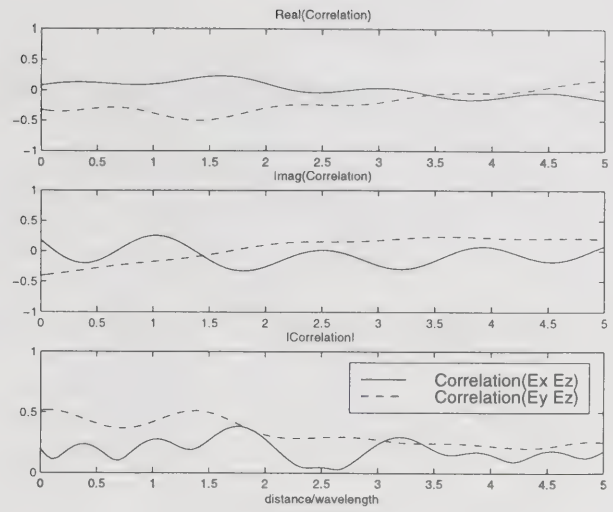


Fig. 6: Correlation plots for the antenna and polarisation diversity case in function of the distance.

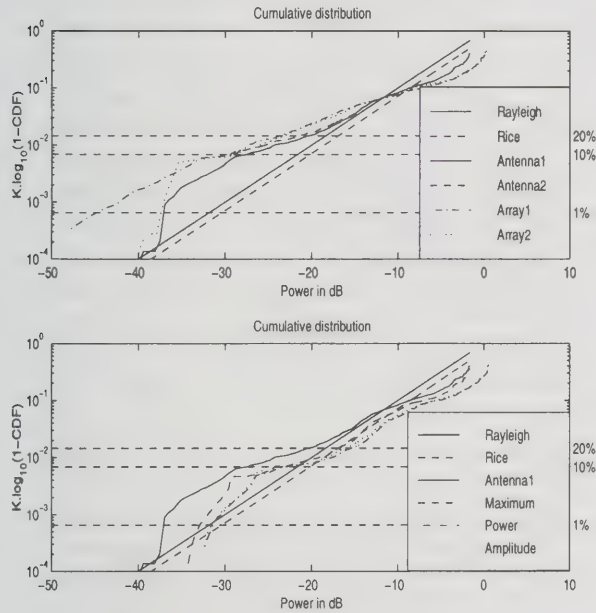


Fig. 5: Cumulative distribution functions of the polarisation diversity.

The third combining method might make use of both previous methods. The second antenna is perpendicular to the first, but at a distance from the first. The correlation plots (Fig. 6) learn us that the polarisation diversity case was already decorrelated, but the separation has also be taken 1 wavelength, like in the 2-antenna diversity case. The CDF's (Fig. 7) show us that the improvement is better than for the case of the polarisation diversity, but still not as good as the 2-antenna case.

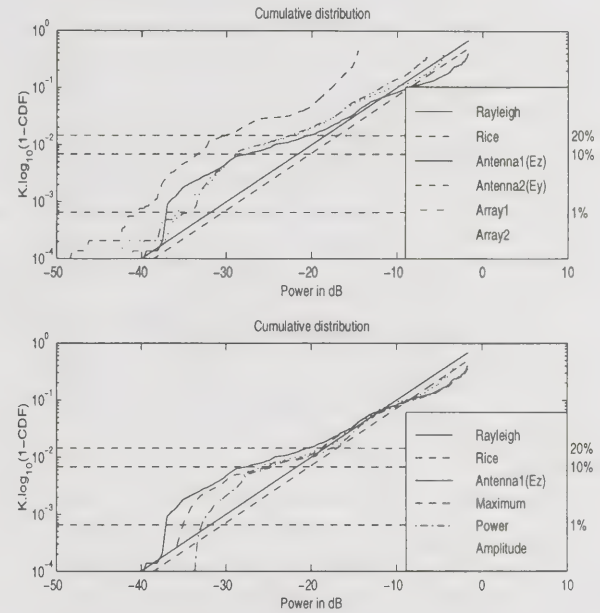


Fig. 7: Cumulative distribution functions of the combined antenna and polarisation diversity.

The 3-antenna case was giving of course nearly identical correlations as for the 2-antenna case, but with the correlation of the elements 1 and 3 being on a distance scale half of that of the 12 and 23 elements. Therefor, it is not shown here. The improvements are better than for the case of 2 antennas (of the order of 3 dB, Fig. 8). This is not sufficient to justify the added complexity of the receiver.

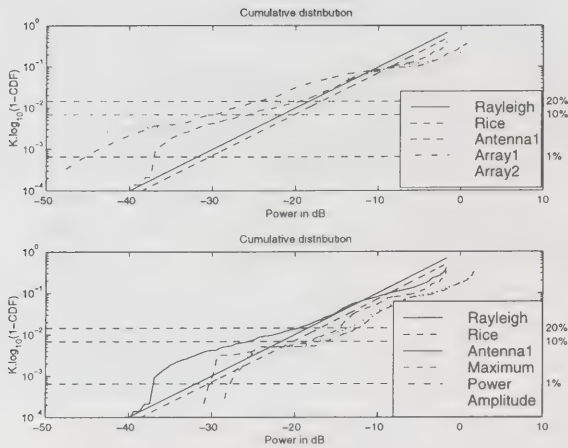


Fig. 8: Cumulative distribution functions of the 3-antenna diversity.

Table 2: Improvement with respect to 1 antenna in 99% of the cases.

	Linear(1)	Circular(2)	Maximum	Amplitude	Power
2-ant.	-7.0	-0.1	4.3	5.5	5.7
Pol.	-22.0	-0.3	0.4	3.0	3.0
a&p	2.8	1.4	2.0	4.2	4.2
3-ant.	-7.5	-8.7	6.8	9.0	9.3

Table 3: Improvement with respect to 1 antenna in 90% of the cases.

	Linear(1)	Circular(2)	Maximum	Amplitude	Power
2-ant.	-1.6	-1.4	3.2	4.1	4.5
Pol.	-2.4	-1.3	0.9	1.7	1.8
a&p	-1.4	-1.0	2.3	1.7	2.9
3-ant.	-4.4	-2.0	7.5	8.2	9.2

Conclusions

We can conclude that in all cases the power combining method is the best, followed closely with amplitude and maximum combinations, and that antenna diversity is better than polarisation diversity. The figures of Table 2 may look high, but they are the effective improvements at system level. Indeed an improvement of the signal at places, where it is already acceptable, does not improve the coverage significantly, but it does where the level is sufficiently low.

Acknowledgement

This research has been supported by the IT Programme of the Flemish Government (ISIS).

All results are summarised in Table 2 for the 99% level and in Table 3 for the 90% level, where the improvements of the signal levels with respect to the single antenna case for the 5 combination methods and the 4 diversity cases are recapitulated. As one can see from the CDF's, the improvements at 99% will be more significant, while the levels at 90% will be less impressive.

References

- [1] Rodney G. Vaughan & J. Bach Andersen, Antenna Diversity in Mobile Communications, IEEE trans. on vehicular technology, Vol. VT-36, no. 4, November 1987
- [2] Iris De Coster, G. Anon Madariaga, B. Pazos Souto, E. Van Lil & F. Pérez-Fontán, An Extended Propagation Software Package For Indoor Communication Systems, Proc. ICAP 1997, Edinburgh, UK, April 14-17, 1997, p.2.298-p.2.301
- [3] Martin Stepler, ComNets Annual Report 1996, RWTH-Aachen
- [4] Silvia Ruiz-Boqué & Ramón Agustí Comes, Polarisation diversity for indoor mobile communications, COST 231 TD(93)
- [5] ITU-R Recomm. nr. PN.1057, 1994 ITU; Radiocomm. sector, 1211 Geneva 20, Switzerland

Spatio-Temporal Equalization for Wireless Communication

Joshua Alspector¹, John D. Norgard, and John Parish
Department of Electrical and Computer Engineering
University of Colorado at Colorado Springs
Colorado Springs, CO 80933-7150 USA

Abstract. Problems with standard cellular systems include limited bandwidth, susceptibility to interference and limits on the number of users per cell. Part of the bandwidth problem can be improved by adding adaptive equalization to standard wireless digital data links. Interference from other users and limitations on the number of users per cell are problems that can be ameliorated by adaptive beamforming antenna arrays. Both spatial (beamforming) and temporal (equalization) processing can be optimized jointly, thus providing better performance than either separately.

Adaptive Non-Linear Equalization

Problems with standard cellular systems include limited bandwidth, susceptibility to interference and limits on the number of users per cell. Part of the bandwidth problem can be ameliorated by adding adaptive equalization to standard wireless digital data links. It is well known that this technique can reduce the bit-error-rate (BER) and increase the bandwidth of a wireless link. However, providing for equalization adds system cost and power in this highly competitive business. Furthermore, in a digital data system with repeaters, like a wireless local-area-network (WLAN), the equalization problem is complicated by the non-linearity of the channel.

We have previously demonstrated how adaptive non-linear equalization can be accomplished by an adaptive learning microchip [Alspector92, 96] which can potentially have power dissipation about a factor of 100 less than using standard digital signal processor (DSP) chips for the same task [Jayakumar93, 94]. The chip uses analog computation and neural network learning ideas which results in a higher computational density and lower power dissipation than standard digital approaches. Overall, the training time per pattern on-chip is quite similar to our simulations. However, in real-time, it is about 1,000 times as fast for a single chip and is even faster for multiple chip systems. The speed for either learning or evaluation is roughly 100 million connections per second per chip.

The idea is to use a known bit sequence to train the equalizer weights to clean up and recover the original signal from the noisy and distorted received signal. The standard linear equalizer is a single neuron network. By using multiple neurons with a hidden layer, one can also perform non-linear equalization which will be important for WLANs. A typical communications channel can be characterized [Proakis89] in the baseband by

$$r_k = h_k s_0 + \sum_{j \neq k, j=-p}^q h_j s_{k-j} + n_k,$$

where r_k is the received signal at the sampling instant, $\{s_j\}$ is the sequence of transmitted

symbols, $\{h_j\}$ is the impulse response of the channel, n_k is the additive noise at the sampling period (assumed to be white Gaussian with variance of σ^2) and $m = p + q$ is the channel memory. The second term on the right hand side of equation (1) is the Inter-Symbol Interference (ISI) caused by the dispersive effects of the channel. A conventional adaptive linear equalizer consists of a tapped delay line FIR filter whose output \hat{s}_{k-d} is an estimate of the symbol d sampling periods ago:

$\hat{s}_{k-d} = W_k^T \bullet \{r_k\}$. An adaptive algorithm tries to minimize an error measure by adjusting the tap weight vector:

$W_k = \{w_0, w_1, \dots, w_{m-1}\}_k$. Inspection of this equation reveals a structure that is similar to a

¹ JA: josh@eas.uccs.edu, (719) 262-3510,
fax: (719) 262-3589

one neuron perceptron with the tapped delay line at the input layer and the summation and thresholding performed at the neuron. This is like any signal processing operation with 1 bit digital restoration. This can be performed in parallel, analog, low-power, compact hardware.

Analog Signal Processing Implementations

Although equalization is usually performed with digital signal processors (DSPs), a higher computational density can be realized by using analog computation, especially for the common signal processing operation of multiply-accumulate. Standard DSPs are expensive and power hungry and we propose to replace them with power-efficient analog neural-style processors. The multiply-accumulate operation in analog electronics can be performed as illustrated by Figure 1 which shows an electronic implementation for a vector-matrix multiplier. The matrix elements are

conductances or weights, w_{ji} , implemented as variable resistors. The weight layer processes an input vector \bar{x} to create an output vector \bar{y} .

The inputs to the layer (the output of the previous layer) are voltages. After passing through the resistors, the voltages become currents. The currents are then easily summed along the input line to the neurons or threshold gates. Thus, the net input to the neurons, \bar{v} , is a vector whose components are $v_j = \sum_i w_{ji} x_i$.

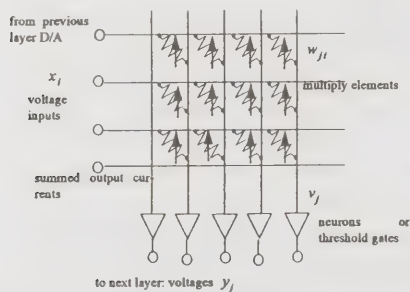


Figure 1. A simple multiply-accumulate processor.

An analog multiply accumulate processor trades off high levels of numerical precision and flexibility for a far higher computational density when compared to a standard DSP. In DSPs, a fast, precision digital multiplication requires a large number of transistors. Thus, a digital implementation can have high precision and

flexibility but at the expense of size on a chip and correspondingly higher power. Another disadvantage of digital techniques is that there is no graceful interface with the natural analog world. In the case of equalization, the sampled analog signals are the interface which would not require A/D conversion.

Adaptive Antenna Arrays

Interference from other users and limitations on the number of users per cell are problems that can be ameliorated by adaptive antenna arrays. Beamforming arrays at the base station can divide the cell into sectors thereby reducing interference and increasing the potential number of users by reusing the frequency within the cell. Information from the beamforming arrays at the base station can be relayed to higher levels of processing for the purposes of location management.

The idea of adaptive learning in VLSI which was demonstrated for equalization can be extended to include finding the maximum of signal strength in beamforming antenna arrays. This can be done without power limitations at the base station but an additional benefit accrues if a lightweight, low power method of beamforming can be applied at the mobile terminal. The same adaptive learning ideas used for equalization can be applied to adaptive beamforming. In equalization, the target function for learning is the actual bit stream sent (which is a fixed pattern in the midamble of the GSM system). In beamforming, maximum signal strength or maximum clarity ("eye" opening between a "0" and a "1") can be used. Fast, low-power processing will be needed to do this. Because a factor of 100 fewer transistors are needed for the multiply-accumulate function compared to a digital chip, an analog VLSI chip will have greater computational density and lower power dissipation than a DSP.

A planar phased array of similar antenna elements can be used to form a directive beam which can be steered to any desired direction in the upper hemisphere above the array by suitably adjusting the phase shifters to the elements. The planar array can be constructed from M linear arrays of N elements each lying in the xy plane with the normal to the array in the z-direction. It is assumed that the antennas

are uniformly spaced with M elements along the x -axis (with an equal-distant spacing of d_x) and with N elements along the y -axis (with an equal-distant spacing of d_y), forming a rectangular grid. The directivity of the array increases with the total number of elements MN .

The complex intensity (magnitude and phase) of the excitation of the mn array element is of the form $I_{mn} = M_{mn} e^{j\varphi_{mn}}$. The complex excitation of the array is assumed to be uniform in magnitude and progressive in phase, i.e., the magnitude and phase of the excitation intensity of the mn phase shifter is of the form $M_{mn} = M_0$, a fixed magnitude and $\varphi_{mn} = (m-1)\varphi_{x0} + (n-1)\varphi_{y0}$. To steer the beam to the θ_0, φ_0 direction, the phase shifters must be set to $\varphi_{x0} = -kd_x \sin \theta_0 \cos \varphi_0$ and $\varphi_{y0} = -kd_y \sin \theta_0 \sin \varphi_0$ where k is the wave number.

Neural Network Spatial Processor

This section describes a neural network spatial processor. A computer simulation demonstrates potential for the neural network to improve antenna array processing efficiency and increase performance in specific situations. As discussed previously, antenna arrays used in wireless communications often incorporate super-resolution algorithms that improve detection performance. Super-resolution algorithms can not only provide precise direction of arrival (DOA), but can also separate signals that the array receives simultaneously at the same frequency. Signals that the array receives simultaneously at the same frequency will be called *cochannel* signals in this section, and simulations for this section will use a narrow band approximation for cochannel signal characteristics. The simulation discussed below predicts that a Radial Basis Function Neural Network (RBFNN) will perform robustly for the case of a uniform linear array (ULA) in the presence of a correlated cochannel signal at a given DOA. The RBFNN is shown to perform as well or better than a widely used super-resolution algorithm known as the Multiple Signal Classification (MUSIC) algorithm.

Simulated Array. The ULA in the simulation has six elements .25 m apart, and the array operates at 400 MHz, giving $\lambda / 3$ spacing.

Neural Network. The RBFNN has three layers. The input layer performs a nonlinear mapping from the signal space to the array response feature space. The I 'th component in the feature space is the output of the I 'th antenna element. The antenna system preprocesses the data by calculating the product $I_{mn} I_{mn}^H$, and then the off-diagonal matrix elements, arranged as vectors, are the inputs to the neural network shown in Figure 2. The inputs are normalized such that the total power in each vector is 1. Since the spatial correlation matrix is complex, the real and imaginary parts of each element are included as individual real inputs to the neural network. The network is trained using a fixed number of radial basis functions, and the LMS technique (generalized inverse) is applied.

Simulations. Two different simulations are reported. In the first case, the MUSIC algorithm is used to track a moving source in the presence of a fixed, correlated cochannel signal. In the second case, the RBFNN performs the same operations. The results of the two different methods are compared. The signal power matrix for a correlation coefficient is

$$P = \begin{pmatrix} p & \gamma^* \\ \gamma p & |\gamma|^2 \end{pmatrix}, \text{ where } p \text{ is the signal power.}$$

The correlated signal was simulated by multiplying the first signal by γ and then changing the DOA. The RBFNN is trained using a signal contaminated by Gaussian noise. For this simulation, the SNR of the first signal was chosen as 20 dB, and the SNR of the correlated signal was 10 dB.

The simulation compares the performance of the two algorithms accurately tracking a source through a range of DOA. For perfect tracking, the lines should coincide in Figure 3, but the results show that the accuracy decreases as the separation between the sources increases. When two correlated sources are used, the MUSIC algorithm reveals only one DOA, and the sensed DOA is biased toward the correlated cochannel source. The simulation was performed again using the RBFNN, as shown in Figure 4. The

RBFNN does not have the bias that was demonstrated for the MUSIC algorithm. The RBFNN method achieves a good result without requiring the eigenvector computation as in the MUSIC algorithm. The RBFNN results were achieved using only 22 training vectors, whereas the MUSIC algorithm required 88 samples of the array manifold.

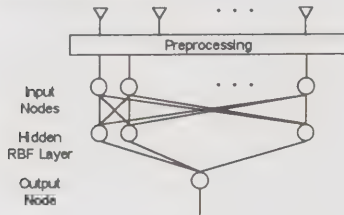


Figure 2. Neural network operates on spatial correlation matrix.

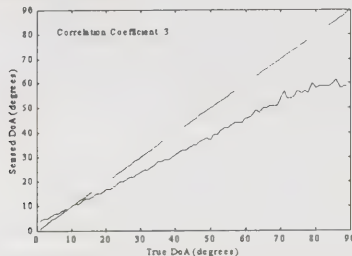


Figure 3. MUSIC algorithm shows increased bias for increased correlation.

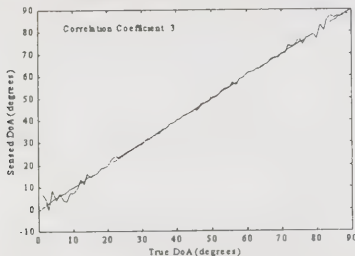


Figure 4. RBFNN tracking performance for increase correlation.

Joint Adaptive Spatial and Temporal Signal Optimization

It is important to note that the adaptive technique allows for joint optimization of both the beamforming direction (or weighted directions) and the equalization function (the weighted sum of sampled delays). This means that both spatial (beamforming) and temporal (equalization) processing can be optimized jointly, thus providing better performance than either separately. This is important if, for example, a particular multipath direction should be rejected to optimally improve the signal to noise ratio. Optimizing on signal strength alone

will not provide the cleanest signal. Rather, a narrowly focused beam on the cleanest signal will do better. If there are several clean signals but they are weak, a weighted equalizer in combination with selective beamforming will do a better job than either technique alone or both techniques optimized separately. A custom, analog VLSI chip based on principles demonstrated above can do either or both jobs in a portable, low-power, compact, low-cost implementation.

References

- [Alspector92] J. Alspector, A. Jayakumar, and S. Luna, "Experimental Evaluation of Learning in a Neural Microsystem," in *Neural Information Processing Systems 4*, ed. J.E. Moody, S.J. Hanson, R.P. Lippmann, Morgan Kaufman, San Mateo, CA, 1992, pp. 871-878.
- [Alspector96] Alspector, J., T.X Brown, and A. Jayakumar, "Adaptive equalizer using self learning neural network," U.S. Patent #: 5,504,780 issued: April 2, 1996.
- [Jayakumar93] A. Jayakumar, J. Alspector, "An Analog Neural Network Co-Processor System for Rapid Prototyping of Telecommunications Applications", at the IEEE/INNS International Workshop on Application of Neural Networks to Telecommunications, October 18-20, 1993, Princeton, New Jersey and in the Proceedings of the International Workshop on Application of Neural Networks to Telecommunications, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1993, pp. 13-19.
- [Jayakumar94] A. Jayakumar, and J. Alspector, "A Neural Network based Adaptive Equalizer for Digital Mobile Radio", in *Government Microcircuit Applications Conference, 1994 Digest of Papers*, (San Diego, CA, November, 1994), pp. 189-190.
- [Proakis89] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1989.
- [Southall95], H. Southall, J. Simmers, and T. O'Donnell, "Direction Finding in Phased Arrays With a Neural Network Beamformer," *IEEE Trans. on Antennas and Propagation*, vol. 43, no. 12, December 1995.
- [Zooghby97] A.H. El Zooghby, C.G. Christadoulou, M. Georgiopoulos, "Neural Network Approach for Direction of Arrival Estimation," *Proc. of the SPIE*, vol. 3077, pp. 572-582, April, 1997.

Interference Resistant Modulation Using Transform Domain Processing

Chad S. Bergstrom and J. Scott Chuprun

Motorola Space and Systems Technology Group
8201 E. McDowell Road, MD H1175
Scottsdale, Arizona, USA 85257-1417
602-441-3278, 602-441-2584 (fax), p21191@email.mot.com

Abstract

Digital mobile battlefield environments impose challenging obstacles to battle group communications and data management. As future communication capabilities cascade from tactical commanders to individual soldiers, spectral clutter and hostile jamming complicate the task of preserving data throughput. Frequency hop (FH) and direct sequence (DS) methodologies have inherent strengths that must be exploited to overcome communication obstacles. In this paper, we present a hybrid methodology that combines FH waveforms with pulse concealment and transform domain signal processing in order to overcome interference, hostile detection, and jamming. Rejection algorithms are demonstrated which significantly improve system performance in the presence of high-power partial-band jammers, co-site communication signals, and fading.

Introduction

Recent studies have shown that co-site and co-channel interference pose especially formidable obstacles to effective battlefield communications[1]. In order to combat ambient and hostile interference, transform domain signal processing may be applied within a spread spectrum communications architecture. Programmable radios can use nonlinear signal processing methods in concert with FH and DS waveforms to mitigate interference while preserving high data throughput, even in the absence of coding.

The properties of DS waveforms inherently provide for interference rejection, low probability of intercept (LPI) and low probability of detection (LPD). However, practical power control limitations are well known to restrict channel capacity[2][3]. In contrast, FH systems are less dependent on power control, yet can be easier to detect, locate, and jam under certain conditions. Higher hop rates on the order of 1-2 Khops/sec enhance FH performance relative to jammer and detection threats [4][5]. However, DoD requirements for future communication systems include provisions for detection and direction of arrival (DOA) capability[6]. Hence, any communication methodology used by U.S. and allied forces must be able to defeat similar detection challenges posed by enemy battle units. In order to meet the challenges posed by the future digital battlefield, we present a hybrid FH/DS approach that preserves LPI/LPD characteristics and tolerates severe interference environments.

Approach

Figure 1 illustrates an example FH/DS hybrid spectrum, including typical interference sources. In this paper, we focus on adaptive mitigation of partial band interference and jammers, in contrast to research which concentrates primarily on the narrow band case [7]. In this paper, both time domain and frequency domain techniques are used for state estimation and interference mitigation, including spectrum estimation and inverse filter countermeasures, and transform domain adaptive inverse weight countermeasures. Prior-art methods, such as frequency domain magnitude normalization, can improve system performance in interference conditions, but impose BER degradation when interference power is low. In contrast, the methods considered in this paper closely follow the theoretical system performance in the absence of jammers. Hence, the methods presented here meet a "do no harm" prerequisite, while significantly improving performance over prior art normalization techniques. Referring again to Figure 1, each hop frequency f_n is spread over a corresponding channel consisting of bandwidth W_{sn} , for a total spreading bandwidth of W_t . When operating near or below the ambient noise floor, spreading the hop impulse in this manner results in a spectral distribution that is less susceptible to follower-jammer and detection threats, while allowing for the application of anti-jam algorithms. Note that when the DS processing gain is low, the hybrid waveform will approach the pure FH system. Since spreading serves to conceal the hop impulse, slower hop rates may be implemented which are less than the current state-of-

the-art. These lower hop rates will provide improved battery life for man-pack or hand-held radios. Furthermore, since DSPN is applied primarily for the purposes of impulse concealment and anti-jam functionality, power control issues are less problematic. Hence, a small number of spreading sequences may be used primarily to prevent the damaging effects of hop collisions between sub-nets.

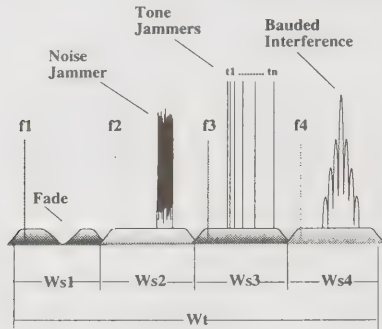


Fig. 1 - Frequency Hop / DS spectrum and interference sources, including Noise Jammers, Tone Jammers, Bauded Interference, and Fading.

In the analysis used for this paper, a QPSK source is used for the I,Q data stream with a DS processing gain of 24 dB. Channel interference sources include a frequency selective fading component, $H(s)$, occurring over the symbol duration. For wideband interference sources and shadowing fades, channel coding can provide sufficient gain to close the link, albeit at a reduction in data throughput. In order to evaluate the effectiveness of preprocessing functions in preserving throughput for the hybrid FH/DS system, we consider such functions in the absence of channel coding.

At the receiver, each W_{sn} channel is tuned corresponding to the Hop Sequence, with timing given by t_{ref} . Advances in accurate, stable oscillators make near-synchronous FH systems possible for near term communication systems [8]. Following tuning, corrupted channel data is passed to an Adaptation Processor, which performs periodic State Estimation, stores the state information to Interference State memory I, and, in parallel, performs Interference Countermeasures corresponding to the prior Interference State. In order to perform these estimation and countermeasure tasks, emerging device technologies may be applied which enable advanced signal processing functions within a programmable radio architecture[9][10][11].

Countermeasures

Spectrum Estimate Countermeasure For the hybrid FH/DS system, the individual channel characteristics for

each W_{sn} are similar to white noise. As such, perturbations in the channel serve to color or distort the spectral characteristic from this ideal. Hence, by obtaining an estimate of the spectral envelope, and by using this estimate to inverse filter the corresponding data, a whitened, or "restored" spectrum may be obtained. In order to ensure that the underlying phase information is undisturbed, an all-pole spectral envelope estimate is required. Equation 1 illustrates an all-pole spectral model, where a_k represents the prediction coefficients for an Nth order model that is obtained using classic autocorrelation and durbin recursion [12].

$$H(z) = \frac{1}{1 + \sum_{k=1}^N a_k z^{-k}} \quad \text{Eq. 1}$$

Once the a_k prediction coefficients are computed, the inverse whitening filter is given by Equation 2, which provides the residual waveform. Hence, Eq. 2 represents the countermeasure that is applied to the received data $x(n)$ by the Adaptation Processor in order to whiten the corrupted spectrum.

$$y(n) = x(n) + \sum_{k=1}^N a(k)x(n-k) \quad \text{Eq. 2}$$

In order to illustrate the inverse filter approach, a partial band noise jammer having bandwidth $B = (0.25)W_{sn}$ is applied to the hybrid FH/DS signal within the hop bandwidth with a channel SNR of -15 dB. Using a reasonable number of prediction coefficients, for example $N = 24$, the inverse filter is applied using Eq. 2. Without reduction in data throughput, the resulting spectral content is whitened, restoring data that had been lost to jamming. To graphically illustrate the effect of the inverse filter countermeasure, Figure 2 shows the effect of the partial band noise jammer on the data QPSK constellation before and after the inverse filter mitigation. Note that the I,Q state information is obliterated by the damaging effects of the high power jammer. Following application of the inverse filter countermeasure, the I,Q state information is restored.

In addition to the damaging effects of narrow band jammers, partial band jammers, and cosite signal interference, communication link integrity can suffer from shadowing and frequency selective fading. Figure 3 illustrates the BER improvement that is obtained using the inverse filter countermeasure in the presence of frequency selective fading. Note that the damaging effects of the fade are largely removed by the inverse filter

countermeasure over a range of channel noise levels. Resulting performance is restored to near fade free levels.

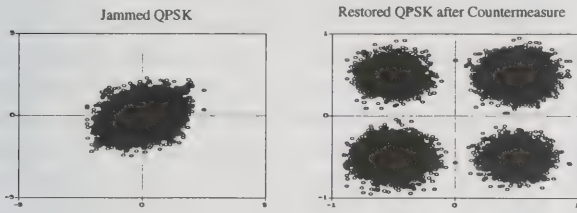


Fig. 2 - QPSK Constellation in presence of high-power partial band noise jammer, showing destruction of I,Q state information, and restored QPSK constellation after application of inverse filter countermeasure.

In high power interference, the inverse filter countermeasure provides improvement in BER performance at high J/S levels, while approaching the asymptotic limit at low J/S levels. Hence, in contrast to other techniques which can induce significant errors in relatively clean conditions, application of this method at low or negligible jammer power does not induce undue BER performance distortion. In comparison, the adaptive inverse weight countermeasure described below has similar asymptotic behavior, yet provides superior BER improvement than the inverse filter for high-power interference.

Adaptive Inverse Weight Countermeasure The second countermeasure method considered for this paper involves transform domain signal processing for interference mitigation. Selective spectral limiting has long proven effective for jammer and interference mitigation [13]. We improve upon such approaches by computing modal spectral statistic estimates μ_i , σ_i for each signal component i . These non-arithmetic statistics may be used to selectively target only those statistical outliers that exceed some value $\alpha = \mu + k\sigma$ of the fundamental transmission mode. Figure 4 illustrates a typical multi-mode statistic for the FH/DS signal in the presence of a partial band noise jammer. Using the transform domain magnitude, a fully adaptive interference rejection function is applied that measures the modal statistics, determines the fundamental mode, and applies an interference rejection weight ω that is inversely proportional to the observed deviation δ relative to the dominant mode. In this approach the statistical outliers are reduced by the weight given by Eq. 3, where m is an integer. In addition to statistical outliers, application of Eq. 3 to $\pm n$ proximity samples provides further performance im-

provement by enhancing the attenuation of the interference envelope.

$$\omega = \alpha / [\alpha + [\delta - \alpha]^m] \quad \text{Eq. 3}$$

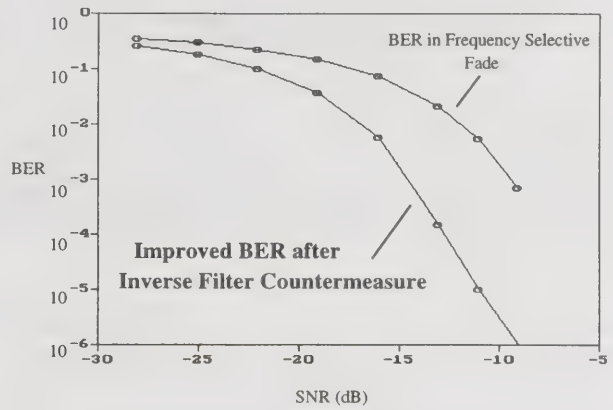


Fig. 3 - BER improvement in frequency selective fade environment after application of Inverse Filter Countermeasure.

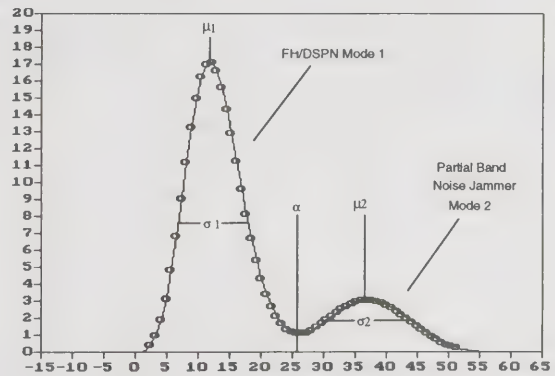


Fig. 4 - Modal statistic estimate for FH/DS signal corrupted by partial band noise jammer.

Figure 5 shows the BER improvement gained by the adaptive weight countermeasure in the presence of QPSK interference encompassing a quarter bandwidth, $(0.25)W_{sn}$, with a channel SNR of -15 dB. Given identical interference bandwidths, the BER performance shown in Figure 5 is similar for QPSK cosite and partial band jammer interference. As with the inverse filter countermeasure, the adaptive inverse weight method approaches the asymptotic limit at low J/S ratios. Note that the adaptive inverse weight dramatically improves BER at very high J/S ratios.

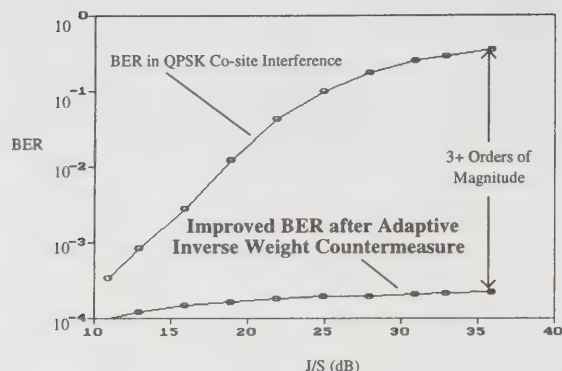


Fig. 5 - BER Improvement in presence of quarter band QPSK interference using Adaptive Inverse Weight countermeasure.

The final interferer considered for the adaptive inverse weighting countermeasure is a four-tone hostile jammer. As with the partial band interferers, the adaptive inverse weight is effective for tone jammers even at very high J/S ratios. This approach is especially beneficial against narrow band interference, which is readily canceled without distortion of the underlying hybrid FH/DS data. Figure 6 shows the BER performance for this method in high-power tone interference. Note that the four tone jammer is mitigated using the adaptive method, resulting in a near-flat BER corresponding to the -15 dB SNR used for this example.

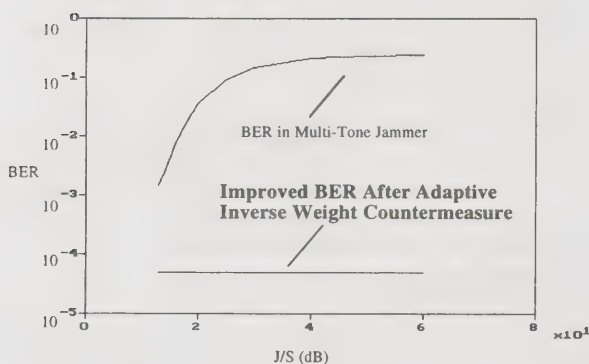


Fig. 6 - BER Improvement in presence of multi-tone jammers using Adaptive Inverse Weight countermeasure.

Conclusion

In this paper, we have presented a hybrid Frequency Hop DSPN transceiver for the digital mobile battlefield. BER improvement has been demonstrated for proposed receiver-based preprocessing functions, in-

cluding an adaptive inverse weight countermeasure based upon spectral modal statistic estimates, and an inverse whitening filter countermeasure based on an all-pole spectral model. Both countermeasures improve BER performance without degrading system throughput, while exhibiting superior performance relative to prior-art full-band normalization methods. Of the two countermeasures, the inverse weight function provides greater BER improvement in the presence of partial band jammers and communication signal interference. The adaptive inverse filter function further provides a mechanism for improving BER performance in the presence of frequency selective fades. Hence, a composite jam/fade scenario will benefit from a cascaded mitigation approach using both methods.

References

- [1] Joe C. Capps, *TXFFI Lessons Learned Report*, July 1, 1997.
- [2] R. Cameron and B. Woerner, "Performance Analysis of CDMA with Imperfect Power Control", *IEEE Transactions on Communications*, Vol. 44, July 1996.
- [3] D. Torrieri, "Frequency Hopping and Future Army Mobile Communications", *Proceedings, ATIRP Annual Conference*, University of Maryland, 1997.
- [4] D. Torrieri, "Fundamental Limitations on Repeater Jamming of Frequency-Hopping Communications", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 4, May, 1989.
- [5] U.S. Government, "Contractor's Design Handbook for Protecting RF Transmissions", Revision 1, July 1994.
- [6] Joseph Mitola III, "Spectrum Supremacy Program Concept", *Mitre Technical Report*, Jan. 7, 1997.
- [7] Kumpumaki, Isohookana, Juntti, "Narrow-Band Interference Rejection Using Transform Domain Signal Processing in a Hybrid FH/DS Spread Spectrum System", *Proceedings, Milcom '97*, Monterey CA, November 1997.
- [8] B. Kaspar, LtCol, USAF, "SUO SAS Phase II Industry Brief", January 29, 1998.
- [9] Analog Devices, "White Paper on the Texas Instruments TMS320C6x"
- [10] Susan Gilfeather and Scott Chuprun, "Complex Arithmetic Processor Performance Metrics on LPI Waveforms", *MILCOM '95 Classified Proceedings*, November, 1995.
- [11] Motorola, "SPEAKeasy II - An IPT Approach to Software Programmable Radio Development", *Proceedings, Milcom '97*, Monterey CA, November 1997.
- [12] Deller, Proakis and Hansen, *Discrete Time Processing of Speech Signals*, p. 297, Macmillan Publishing Co., 1993.
- [13] Simon, Omura, Scholtz, Levitt, *Spread Spectrum Communications Handbook*, p. 451, McGraw-Hill, 1994.

Smart Channel Assignment Algorithm for SDMA Systems

Flavio Piolini, Anna Rolando

CSELT - Centro Studi e Laboratori Telecomunicazioni S.p.A.

Via G. Reiss Romoli 274 -10148 TORINO - ITALY

F. Piolini : tel.: + 39/11/228-7423, fax: +39/11/228-5581, e-mail : flavio.piolini@cse.lt.it

A. Rolando : tel.: + 39/11/228-7531, fax: +39/11/228-7078, e-mail : anna.rolando@cse.lt.it

Abstract

In order to exploit the potentialities of Space Division Multiple Access (SDMA) in mobile cellular systems the conventional radio resource management functions must be revised, especially the channel assignment algorithms. In the paper an algorithm designed to this purpose is presented and analysed in reference scenarios and under various traffic load conditions.

Key words: SDMA, channel assignment algorithm, intra-cell handover

1. Introduction

Mobile cellular systems are characterised by an impressive subscriber increase. In order to satisfy such a large demand, operators have to cope with the scarceness of the limited frequency spectrum. One approach which looks very appealing is the use of adaptive antennas at Base Station site, combined with Space Division Multiple Access (SDMA) technique [1]. SDMA allows the simultaneous use of any conventional channel (frequency, time slot or code) by multiple users of the same cell by exploiting the angular separation among them. This is achieved by taking advantage of the ability of the adaptive antenna technology to distinguish between desired and interfering signals as well as to calculate their direction of arrival. Following the changes in users' locations, the antenna system continuously adapts its radiation pattern, so that the target user is tracked with the main lobe, while nulls of the antenna gain diagram are set in the direction of the interferers. As result the spectral efficiency and hence the capacity of the system increases [2]. On the other hand, SDMA may introduce some critical aspects in the radio resource management. Particularly, the conventional channel assignment algorithms, both in case of call set-up and handover, must be harmonised with the characteristic of the SDMA approach [3,4].

The paper presents a channel assignment algorithm for a generic frequency/time division multiple access (FDMA/TDMA) system applying SDMA and evaluates its performance in reference scenarios and under various traffic load conditions.

2. Channel assignment algorithm

In order to exploit the potentialities of SDMA, the Base Station (BS) must keep track of the position and mobility behaviour of all mobile users (MSs) within the cell. When a user requires a channel, the channel assignment algorithm should select it so that the probability of two co-channel users coming close to each other (and then asking for intracell handover) is minimised. A number of factors can affect the optimal choice; amongst them, the most significant are:

- Cell topography;
- MSs positions;
- MSs mobility characteristics (speed and direction).

The proposed algorithm takes into account these parameters, leading to an optimised allocation of the radio resources. The basic principles the algorithm relies on are described in the sequel.

2.1. Basic principles

First, the area of the cell is subdivided into N sectors (see Figure 1). A regular, rectangular grid is adopted here, nevertheless other partitioning schemes (e.g. specifically tailored on the cell topography) could be used. MSs position within the cell is identified by means of the correspondent sector. Then, for each pair of sectors S_i and S_j belonging to the cell a coefficient is defined as:

$$C_{i,j} = f(P(S_i, S_j), T(S_i, S_j))$$

where:

$P(S_i, S_j)$: is the estimated probability that a mobile entering sector S_i will come to collide with a mobile that at the same instant is located in sector S_j (i.e. the probability that the main antenna beams formed towards the two mobiles will overlap, thus causing an intracell handover);

$T(S_i, S_j)$: is the estimated elapse of time before the collision between a mobile entering sector S_i and a mobile that at the same instant is located in S_j occurs.

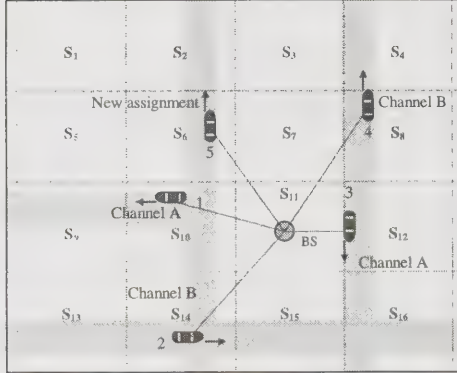


Figure 1: Channel assignment

The coefficients $C_{i,j}$ are used as a metric to sort the candidate channels for channel assignment purpose. Let consider the case when a channel assignment request is submitted by a user located in a generic sector S_h of the cell. If no free channel is available, the spatial reuse technique has to be applied.

According to the proposed algorithm, for each candidate channel z a cost figure W_z is evaluated, based on the coefficients $C_{h,k}$ referring to all the sectors k where the candidate channel is currently in use:

$$W_z = w(C_{h,k} | \text{channel } z \text{ is in use in sector } k)_{k \in \{1, \dots, N\}}$$

Finally, candidate channels are sorted according to their cost, and the one corresponding to the minimum cost value is selected and assigned to the mobile. In order to clarify how the algorithm works, an example is given with reference to the scenario depicted in Figure 1. Mobile 1 and 3 are sharing channel A, while mobile 2 and 4 are sharing channel B. No other channel is available in the cell. If mobile 5, located in S_6 , performs a channel request, then the following cost coefficients are evaluated for channels A and B respectively:

$$W_A = w(C_{6,10}, C_{6,12})$$

$$W_B = w(C_{6,8}, C_{6,14})$$

If $W_A < W_B$ channel A is assigned to mobile 5, otherwise channel B will be chosen.

2.2. Further enhancements

The procedure described so far takes into account only the cell topography and the mobiles positions. To enhance the performance of the algorithm by making the estimation of the handover probability more accurate, the mobility trends (speed and direction) of the mobile users have to be considered. Hereafter it is explained how to include them in the algorithm.

2.2.1. Influence of movements direction

The direction of the MSs can be taken into account by introducing the direction of arrival into a sector as an additional parameter in the evaluation of coefficients $C_{i,j}$. This means that for each couple of sectors S_i and S_j , there will be a number of coefficients:

$$C_{i,j}^{m,n} = f(P(S_i^m, S_j^n), T(S_i^m, S_j^n))$$

where m and n indicate respectively the direction of arrival in sector S_i and sector S_j .

As our partitioning scheme is a regular rectangular grid, for each sector the direction of arrival can be identified with the four sides of the sector. Thus, for a given couple (i,j) , the number of coefficients $C_{i,j}^{m,n}$ will be 16 and m, n can assume the values N (north), S (south), E (east) and W (west).

Applying the proposed enhancement to the example of Figure 1, the following coefficients are computed:

$$W_A = w(C_{6,10}^{S,N}, C_{6,12}^{S,N})$$

$$W_B = w(C_{6,8}^{S,S}, C_{6,14}^{S,W})$$

2.2.2. Influence of speed

The MSs speeds can be taken into account by defining different speed classes and by determining one coefficient for each speed class. This means that, assuming that the speed classes are M , for each couple of sectors S_i and S_j , there will be M^2 coefficients:

$$C^{p,q}_{i,j} = f(P(S_i^p, S_j^q), T(S_i^p, S_j^q))$$

where p and q indicate respectively the speed class of the MS in sector S_i and of the MS in sector S_j .

2.3. Matrix coefficients computing

The coefficients $C_{i,j}$ are computed during the system initialisation phase. During this phase, the BS monitors the behaviour of each active user served by the cell, keeps track of the ordered list of sectors that any user comes through following its own path within the cell and of the entering times into those sectors. It records the user collisions with other users, the positions (in terms of sectors), where each collision

takes place, as well as the time instants when these events occur.

From these samples the values of $P(S_i, S_j)$ and $T(S_i, S_j)$ are derived, and, finally, the coefficients C_{ij} are drawn. An example of how the procedure works is given in the sequel, with reference to the case depicted in Figure 2.

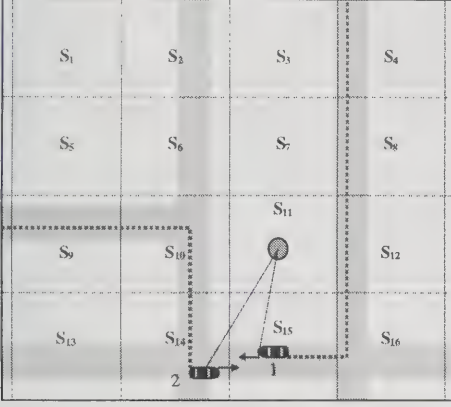


Figure 2: Matrix coefficients computing (1)

Two users, namely user 1 and user 2, are moving through the area of the cell. The path followed by each of them is stored in a dedicated record, which is updated each time the user enters into a new sector. At the time depicted in Figure 2, when the two users come to collide (i.e. their angular separation with respect to the BS goes below a pre-set threshold), the records stored so far for them are the following:

Sector	S4	S8	S12	S16	S15	User
Entering time	t4	t8	t12	t16	t15	1

Sector	S9	S10	S14	User
Entering time	t9	t10	t14	2

Starting from them, a new sample of $T(S_i, S_j)$ is evaluated, for each couple of sector S_i and S_j which has been visited by one of the two user respectively. With reference to Figure 3, the coefficients $T_{14,15}^{N, E}$, $T_{15,10}^{E, W}$, $T_{16,10}^{N, W}$, $T_{10,12}^{W, N}$, $T_{12,9}^{N, W}$ and $T_{9,8}^{W, N}$ are computed.

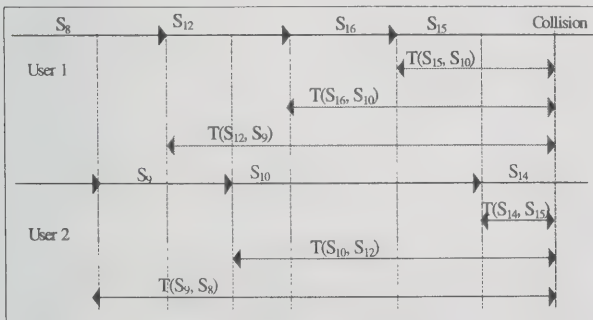


Figure 3: Matrix coefficients computing (2)

At the end of the initialisation phase, the final values of $T^{m,n}(S_i^m, S_j^n)$ are obtained by averaging the values collected for each instance of the random process. The values of $P^{m,n}(S_i^m, S_j^n)$ are determined by averaging the number of collisions of the users entering in S_i with the users that are in S_j respect to the total number of users entering in S_i . Finally, the coefficients $C_{ij}^{m,n}(S_i^m, S_j^n)$ are computed.

3. Simulation tool

A software simulation tool has been developed to evaluate the performance of the channel assignment algorithm. The basic features are described in [3] and summarised hereafter. A single rectangular shaped cell with the Base Station positioned in its centre is considered. Inside the area, streets are traced according to a rectangular grid; a generic settlement can be considered, provided that the roads can run either parallel or perpendicular to each other. The mobile terminals are generated and move onto this grid according to the implemented call and mobility models. The number of ongoing calls inside the area is an input parameter and is kept constant along the whole simulation time. The number of physical traffic channels assigned to the cell is also fixed as an input parameter (15 channels for the simulations). The tool allows to compare the performance of the algorithm under study with respect to a reference one, since initial channel assignment, handover triggering and execution can be performed in parallel according to the two criteria, one independently from the other. The reference algorithm (hereafter indicated as A1) chooses the channel randomly among those (if any) satisfying the basic requirement about the minimum angular separation among co-channel users. The handover is initiated when the angular separation between any co-channel users goes under a value related to the minimum beamwidth of the antenna radiative pattern. In the simulation a fixed value equal to 15° is assumed. When a handover is required, the channel assignment algorithm is applied. If no useful channel is available, the call is dropped and a handover failure is recorded.

The outputs of the simulation are expressed in terms of Intra-cell Handover rate per call and Assignment Failure rate per cell. With reference to the proposed channel assignment algorithm (hereafter referred as A2), the coefficients and cost parameters are evaluated according to the following functions:

$$f(P(S_i, S_j), T(S_i, S_j)) = 1 / P(S_i, S_j) * T(S_i, S_j)$$

$$w(C_{h,k}) = 1 / \sum (1/C_{h,k})$$

Furthermore no speed class subdivision has been adopted.

4. Simulation scenarios and results

Three scenarios have been selected in order to analyse the performance of the algorithm with respect to area topography and mobility features. *Scenario A* (see Figure 4) represents a rural area crossed by highways where only high speed mobiles exist. *Scenario B* (see Figure 5) represents a urban area with low speed mobiles. *Scenario C* (see Figure 6) represents a residential area, where only very low speed mobiles are supposed to be present. The performance of the channel assignment algorithm has been analysed in each reference scenario for different values of the mean spatial reuse factor RF, defined as the ratio between the number of ongoing calls and the number of available traffic channels in the cell (R_{HO}).

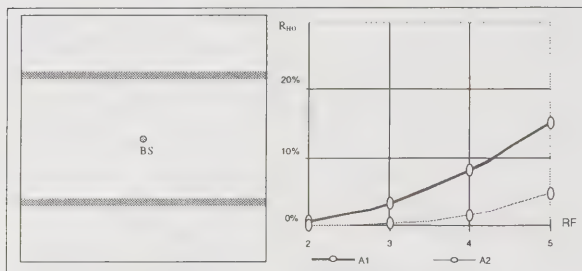


Figure 4: Scenario A

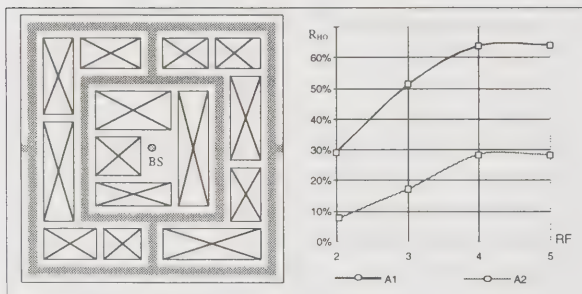


Figure 5: Scenario B

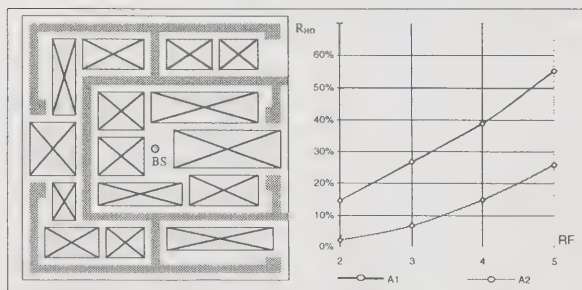


Figure 6: Scenario C

The results (see Figure 4-5-6) show that, by exploiting information about the mobiles spatial distribution in the channel assignment process, the proposed algorithm provides a substantial

improvement in SDMA performance. Note that the characteristics of the environment can heavily influence the performance of each channel assignment algorithm. For instance, in a scenario where the roads disposal is such that the users can move along a low number of well predictable paths, the algorithm is likely to perform well. On the contrary, in a scenario characterised by many alternative ways, the performance of the algorithm worsen. For example, for $RF=3$, the ratio between $(R_{HO})_{A1}$ and $(R_{HO})_{A2}$, is 12,5 in scenario A, while it is reduced to 3.9 in scenario B and C. Concerning the impact of traffic load on the algorithm performance, it results that as the traffic load, and then the reuse factor, increases, the relative gain of algorithms A2 with respect to A1 becomes lower. This can be justified by the fact that, when the users density inside the area served by the cell is very high, it is likely that the interference spatial distribution becomes, on average, similar on different traffic channels, thus making the performance less sensitive to the channel assignment policy.

5. Conclusion

A channel assignment algorithm for SDMA mobile systems has been described and analysed in reference deployment scenarios and under various traffic load conditions by means of software simulation. The results obtained show that the algorithm performance, evaluated in terms of average number of intra-cell handovers per call, are considerably better respect to other algorithms that do not take into account the users spatial distribution and their mobility features.

6. References

- [1] M. Tangemann et alii, "Introducing Adaptive Antenna Concepts in Mobile Communication Systems," Proceedings of RACE Mobile Communications Workshop, Amsterdam, The Netherlands, May 1994, pp. 714-727.
- [2] E. Buracchini, F. Muratore, V. Palestini, M. Sinibaldi, "Performance analysis of a mobile system based on combined SDMA/CDMA access technique," Proceedings of IEEE ISSSTA 96, Mainz, Germany, Sept. 22-25 1996, pp.370-374.
- [3] F. Piolini, A. Rolando, "Channel assignment algorithms for SDMA systems", Proceedings of ICPWC'97, Bombay - India, 17-20 Dec. 1997.
- [4] T. Ohgane, Y. Ogawa, K. Itoh, "A Study on a Channel Allocation Scheme with an Adaptive Array in SDMA," Proceedings of VTC 97, Phoenix, AZ, USA, May 5-7 1997, pp.725-729.

MAI Cancellation By Optimum Single-User Detection For DS/CDMA Systems Using Random Binary Codes

Naresh R. Patel and Timothy O'Farrell.

School of Electronic and Electrical Engineering,
The University of Leeds, Leeds, W. Yorkshire, LS2 9JT, U.K.
Tel: (0113) 233 2098, Fax: (0113) 233 2032
email: eenrnp@electeng.leeds.ac.uk

Abstract

The Optimum Single-user Detector (OSD) for a DS/CDMA based communications system is investigated using random binary signature codes of length 32 and 64 chips. The BER performances are obtained by Monte-Carlo simulation for spectral efficiencies of 50% and 100%. Results show that the OSD performs significantly better than the Matched Filter Detector (MFD) with a modest increase in complexity. It is also seen that noise enhancement is induced due to MAI cancellation which degrades the OSD performance.

Introduction

Optimum detection of a signal that is corrupted by Additive White Gaussian Noise (AWGN) is achieved by the Matched Filter Detector (MFD). The real valued impulse response of the MFD is given by the time reversal of the wanted signal. An important quantity which characterises the statistics of the channel noise is the autocorrelation matrix \mathbf{R}_a . For AWGN, this autocorrelation matrix is just an Identity matrix which is scaled by the variance of the noise process. The MFD is the optimum detector when the autocorrelation matrix of the channel noise is represented by the a scaled Identity matrix.

It is widely known that Multiple Access Interference (MAI) is produced due to the interaction between the code sequences of the simultaneous users. It is usually viewed to be characterised by the crosscorrelation between the code sequences which is contained in the crosscorrelation matrix \mathbf{R}_c . This is the basis of multiuser detection. The autocorrelation matrix succinctly contains this information also. It is seen that the autocorrelation matrix produced by MAI has definite sidelobes. The presence of these sidelobes makes the MFD sub-optimal in the presence of MAI. It is this form of noise that limits the performance and hence the capacity of communications systems based upon DS/CDMA. There is a considerable amount of research being carried out to tackle this problem. Most solutions involve multiuser detection. These schemes involve some form of joint detection by a bank of MFD's followed by some algorithm in deciding the wanted data. The optimum multiuser detector was the first

proposed solution [1]. This scheme optimises the decision process after detection by a bank of MFD's in the maximum likelihood sense. This structure is very complex to implement so other sub-optimal derivatives have been proposed [2-4]. These sacrifice performance for a limited decrease in complexity. The major limiting factor of multiuser detectors is their complexity.

This paper considers an Optimum Single-user Detector (OSD) which uses the information contained in \mathbf{R}_a and obtains optimum tap co-efficients in the sense of maximising the output Signal-to-Noise Ratio (SNR). It has a complexity similar to the MFD, but shows a substantially better Bit Error Rate (BER) performance. The OSD is investigated in terms of the BER vs. E_b/N_0 performance for spectral efficiencies of 50% and 100%. Random codes of length 32 and 64 chips were used. Section 1 introduces a model for a synchronous DS/CDMA system in a non-dispersive channel corrupted by AWGN. Optimum single-user detection is introduced in section 2. It is the SNR that is maximised at the output of the detector which leads to the introduction of the OSD in section 3. Simulation results are presented in section 4 followed by the conclusion in section 5.

1. The DS/CDMA model

Baseband transmission of data K simultaneous users transmitting simultaneously through a channel corrupted by AWGN is modelled. The k -th user's data bit $b_k(t)$ is transmitted at time t , for a duration of T_b . A unique code sequence of length N chips is assigned to each user upon the system. The chip duration is T_c . Both the chips and

bits are bipolar in form and are sequences of unit amplitude rectangular pulses. The k -th user's data bit is modulated by its assigned code sequence $a_k(t)$, hence $T_b=NT_c$ and its transmitted signal $s_k(t)$ is given by is expressed by,

$$s_k(t) = b_k(t)a_k(t) \quad (1)$$

where the average power of each user's signal over one code period is normalised to unity. In this model, each user's transmitted signal during one data bit period is its code sequence or its complement depending upon the value of $b_k(t)$ and one data bit of all the signals are considered to be contained within a single-shot detection process. This superposed received signal $r(t)$ at the input of the receiver is given by,

$$r(t) = \sum_{k=1}^K b_k(t)a_k(t) + n(t) = \sum_{k=1}^K s_k(t) + n(t) \quad (2)$$

where $n(t)$ represents AWGN with a two-sided power spectral density of $N_0/2$, present upon reception. The system is considered in the discrete-time domain after sampling takes place at a rate $1/T_c$ at the receiver. Using a vector notation, each discrete time sample will be placed as an element of this received observation vector \mathbf{r} , which represents the N samples made to recover a complete data bit during the single-shot detection process. By definition, the observation vector is the time reversal of the sample function of the received signal due to the way the samples are shifted in the tapped delay line filter. As an extension to Eq.(2), in the vectorial domain, the received signal after N samples is given by,

$$\mathbf{r} = \sum_{k=1}^K b_k \mathbf{a}_k + \mathbf{n} = \sum_{k=1}^K s_k + \mathbf{n} \quad (3)$$

where \mathbf{a}_k , \mathbf{s}_k and \mathbf{n} are observation vectors of the k -th user's code sequence, the k -th users signal and the AWGN process after sampling respectively. Each data bit period is detected and a decision upon the wanted user's data bit is made.

2. Optimum Single-user Detection

As already mentioned, the autocorrelation matrix of a signal succinctly contains information of the second order statistics of the channel noise and it is seen that MAI produces definite sidelobes to this matrix. These sidelobes make the MFD sub-optimal. The OSD uses this knowledge of the statistics of the channel noise (both MAI and AWGN) to obtain an optimum filter response for the detector. In the detection of the w -th

user's signal during a single-shot detection process, the received signal represented by an N -by-1 observation vector \mathbf{r} (Eq. 3) can be written as,

$$\mathbf{r} = b_w \mathbf{a}_w + \left(\sum_{\substack{k=1 \\ k \neq w}}^K b_k \mathbf{a}_k + \mathbf{n} \right) = \mathbf{r}_s + \mathbf{r}_n \quad (4)$$

where \mathbf{r}_s is the wanted signal part of the received signal and \mathbf{r}_n is the corrupting signal which contains the noise component consisting of MAI plus AWGN. The auto-correlation matrix \mathbf{R}_a is defined by the expectation of the outer product of \mathbf{r}_n with itself and is given by [5],

$$\mathbf{R}_a = \frac{1}{K-1} \sum_{\substack{k=1 \\ k \neq w}}^K \mathbf{a}_k \mathbf{a}_k^T + \sigma^2 \mathbf{I}_{N \times N} \quad (5)$$

where σ^2 is the variance of the AWGN process which has a power spectral density of $N_0/2$ and \mathbf{I} is an N -by- N Identity matrix. The first term of \mathbf{R}_a represents the auto-correlation matrix of MAI and the second term is the auto-correlation matrix of AWGN. The impulse response of the detector (linear transversal filter) is given by the filter coefficients \mathbf{h} . The value of \mathbf{h} that maximises the Signal-to-Noise Ratio (SNR) at the output of the OSD detecting for the w -th user's signal is given by [6],

$$\mathbf{h} = \mathbf{R}_a^{-1} \mathbf{a}_w \quad (6)$$

There is a trade-off between MAI cancellation and optimum detection in AWGN i.e. optimising the detection process in the presence of MAI will change the impulse response of the detector which will produce sub-optimal detection in the presence of AWGN and vice versa. As MAI is the limiting factor, an improvement in performance should be seen due to MAI cancellation. It is also clear that the structure of \mathbf{R}_a will vary depending upon the properties of the code set used. A 'better' set of codes produce less MAI and both the MFD and the OSD will perform better.

3. Results

A baseband, synchronous DS/CDMA system was simulated using random binary code assignments. Code lengths of $N=32$ and 64 chips were used for spectral efficiencies of 50% and 100%. The results of BER vs E_b/N_0 were obtained by Monte-Carlo simulation, for both MFD and OSD implementations. **Figure 1** shows the results obtained when code lengths of $N=32$ and 64 were used for a spectral efficiency of 50%. The plot also shows the single-user bound for comparison.

Firstly, the MFD (dashed lines) suffers greatly due to the presence of MAI with an irreducible BER floor at around 4×10^{-2} . In comparison, the OSD (solid lines) performs significantly better than the MFD for both code lengths. Noise enhancement is induced due to MAI cancellation which is why the curves do not follow the single-user bound. The OSD performs better with the shorter code length for the same spectral efficiency over the range investigated. As MAI increases, more adjustments need to be made to the MFD filter coefficients in order to combat MAI. This means that greater noise enhancement is induced. **Figure 2** shows the results obtained for a spectral efficiency of 100%. Again the MFD suffers from an irreducible BER floor at around 10^{-1} and the OSD significantly improves the BER performance. It is seen that the performance of the OSD is degraded by increasing the spectral efficiency. Again this is due to a greater amount of MAI present which will induce more noise enhancement. Over the same range as before ($E_b/N_0=0\text{dB}$ to 20dB) the OSD performs better with the shorter code length.

4. Conclusion

An investigation of the OSD and MFD is presented for a synchronous, baseband DS/CDMA system where random code sequences of length $N=32$ and 64 were assigned for spectral efficiencies of 50% and 100%. The channel noise consists of MAI and AWGN. The BER performance of the MFD is degraded by the presence of MAI and an irreducible BER floor is seen. There are significant performance gains achieved by using the OSD over the MFD, due to MAI cancellation. The OSD does not suffer from an irreducible BER floor over the range investigated. For the same spectral efficiency, the OSD performs better using the shorter code length. The \mathbf{R}_a matrix is an important parameter which contains two elements, one corresponding to the statistics of MAI and the other term which corresponds to the statistics of AWGN. As the number of simultaneous users K increases, a greater amount of MAI is produced and hence the sidelobe energy present in \mathbf{R}_a (which is only dependent upon MAI) increases which induces greater noise enhancement. The degree of noise enhancement is dependent upon the amount of MAI. This can be seen by realising that the optimum detector in the presence of AWGN is the MFD. Any changes made to the impulse response of the detector will make it sub-optimal. This makes the OSD sub-optimal in the presence of AWGN and noise enhancement will ensue. There is a trade-off between MAI cancellation and noise enhancement. If noise enhancement is reduced, then further improvements in performance could be achieved to bring the OSD curves closer to the single-user bound. This is a subject for

further study. The major advantage of the OSD is that its complexity is substantially less than the multi-user counterpart and a bank of OSD's may act as a direct replacement for the MFD's in multi-user cancellation techniques. The performance of the OSD in radio DS/CDMA systems is also the subject of further investigation.

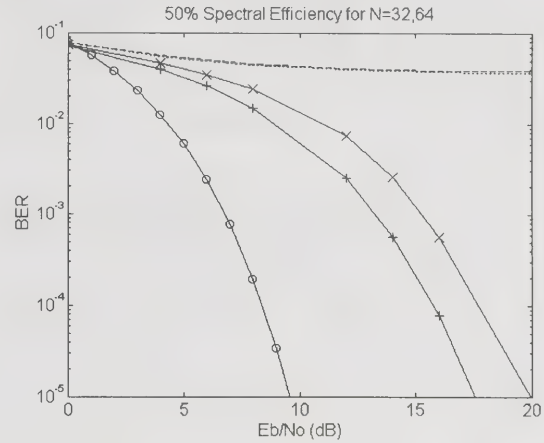


Figure 1: BER vs E_b/N_0 for $N=32$ and 64 for a spectral efficiency of 50%.

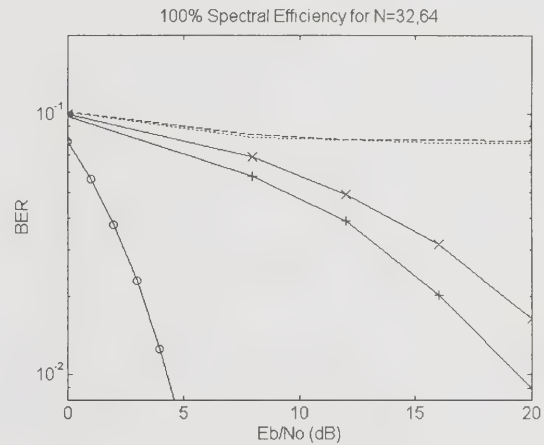


Figure 2: BER vs E_b/N_0 for $N=32$ and 64 for a spectral efficiency of 100%.

Key for the curves:

- single-user bound
- MFD(N=32)
- MFD(N=64)
- +— OSD(N=32)
- x— OSD(N=64)

References

- [1] "Optimum Multi-User Asymptotic Efficiency", S. Verdu, *IEEE Trans. on Commun.*, vol. 34, no. 9, pp 890-897, Sept. 1986.
- [2] "Linear Multi-User Detectors for Synchronous Code-Division Multiple-Access Channels", R. Lupas and S. Verdu, *IEEE Trans. Information Theory*, vol. IT-35, pp. 123-136, Jan. 1989.
- [3] "Near-Optimum Detection in Synchronous Code Division Multiple Access Systems", M.K. Varanasi and B. Aazhang, *IEEE Trans. on Commun.*, vol. 39, no. 5, pp. 725-736, May 1991.
- [4] "A Family Of Suboptimum Detectors for Coherent Multi-User Communications", Z. Xie, R.T. Short and C.K. Rushforth, *IEEE Journal on Selected Areas in Commun.*, vol. 8, no. 4, pp. 683-690, May 1990.
- [5] "Optimum Single-User Detection For DS/CDMA Systems Achieving Complete MAI Cancellation", N.R. Patel and T. O'Farrell, in *Proc. IEEE Int. Symp. on Spread Spectrum Theory and Applications* (Sun City, South Africa, 1998).
- [6] "CDMA Techniques in Optical Fibre Local Area Networks", T. O'Farrell, *PhD. Thesis*, University of Manchester, UK., 1989.

The Application of Low Noise Amplifiers in CDMA Cellular and PCS Systems for Coverage and Capacity Enhancements

Samuel Yang
AirTouch Cellular
3 Park Plaza
Irvine, CA 92614-8535
samuel.yang@airtouch.com

Abstract

The paper examines the use of LNAs as a technique to enhance reverse-link coverage and capacity in a CDMA wireless network. The reduction of system noise figure for a base station is accomplished by installing LNAs between the receive antenna and the radio. Using noise figure theory, the paper develops an analytical model for the SNR improvement. The SNR models are found to be functions of the reverse-link noise rise, which constitutes a measure of reverse-link capacity limits. Since the CDMA system uses reverse-link power control, the improvement in SNR is manifested in the reduction of the mobile's transmit power. In terms of capacity, the reverse-link capacity of a CDMA base station with installed LNAs has not improved. However, the capacity of a neighboring base station (with no LNAs) has improved. This is because the mobile within the coverage area of a base station (with LNA) now transmits less power and contributes less loading to the neighboring base station (with no LNA).

Key words: Code-division multiple access, low noise amplifier, capacity, coverage, CDMA.

Introduction

Over the last few years, cellular and PCS networks using code division multiple access (CDMA) spread-spectrum technology have been widely deployed around the world. Wireless service providers initially deployed these systems to accommodate the increasing traffic demand. However, because of the higher-than-expected demand for many wireless service providers, they are finding the need to exploit their CDMA infrastructure investments to the fullest in terms of coverage and capacity. Recently, low noise amplifiers (LNA) have been used by a number of carriers to augment coverage and capacity of their networks. LNAs, when placed as a front-end in CDMA base stations, provide enhanced range-extension and, in some cases, reverse-link capacity enhancements as well.

Reverse-Link Rise

A typical base station radio front-end is shown in Figure 1 with antenna temperature T_A , cable loss L_1 , radio front-end gain G_2 , and radio front-end NF F_2 . The signal right after the antenna has a signal power of S_{in} , and the SNR at the pre-detection is $(S / N)_{out}$. The composite NF can be written as

$$F_{comp} = F_1 + \frac{F_2 - 1}{G_1} = L_1 + L_1(F_2 - 1) \quad (1)$$

Recall that the composite noise temperature for a cascaded system is [1]

$$T_{comp} = (F_{comp} - 1)T_0 \quad (2)$$

and the system noise temperature is the sum of antenna temperature and composite noise temperature, i.e.,

$$T_{sys} = T_A + T_{comp} \quad (3)$$

If the antenna is aboard a satellite facing the earth, then the antenna temperature T_A would be 290K (i.e., ambient temperature T_0) since the earth's temperature is effectively at 290K. However, in our case, the effectively antenna temperature would be higher than the ambient temperature because the antenna is also receiving interference from other sources. We can use the definition of reverse-link *rise* R to quantify this increase in interference [2]

$$R = \frac{I'_m + I'_t + I'_n + N}{N} \quad (4)$$

where I'_m is the received interference power due to traffic channel transmissions of mobiles in the same cell, I'_t is the received interference power due to traffic channel transmissions of mobiles in other cells, I'_n is the received interference power due to intentional or unintentional jammers, and N is the thermal noise. Since the actual noise seen by the antenna is effectively the numerator of equation (4), the actual noise can be written as

$$\begin{aligned} I'_m + I'_t + I'_n + N &= RN = R(kT_0 W) \\ &= k(RT_0)W = kT_A W \end{aligned} \quad (5)$$

Therefore, the effective antenna temperature (T_A) is equivalent to the rise multiplied by the ambient temperature (RT_0).

Typical Base Station Front-End

The $(S/N)_{out}$ for the system in Figure 1 can be shown to be

$$\begin{aligned} \left(\frac{S}{N}\right)_{out} &= \frac{S_{in}(1/L_1)G_2}{(1/L_1)G_2(kT_{sys}W)} \\ &= \frac{S_{in}(1/L_1)G_2}{(1/L_1)G_2k(T_A + T_{comp})W} = \frac{S_{in}}{k(T_A + T_{comp})W} \end{aligned} \quad (6)$$

where the signal (S_{in}) and the system noise ($kT_{sys}W$) right after the antenna both undergo a loss of L_1 and a gain of G_2 . T_{comp} can be obtained by substituting equation (1) into (2), i.e.,

$$T_{comp} = (F_{comp} - 1)T_0 = (L_1 F_2 - 1)T_0 \quad (7)$$

Substituting (7) into (6) yields

$$\left(\frac{S}{N}\right)_{out} = \frac{S_{in}}{kT_0(R + (L_1 F_2 - 1))W} \quad (8)$$

which can be written as

$$\left(\frac{S}{N}\right)_{out} = \frac{S_{in}}{k\sigma T_0 W} \quad (9)$$

σ is the *noise enhancement* which is given by

$$\sigma = R + (L_1 F_2 - 1) \quad (10)$$

Base Station Front-End with LNA

Figure 2 shows a base station front-end with an LNA between the antenna and the radio. The composite NF of the receiving equipment is

$$F'_{comp} = F'_1 + \frac{F'_2 - 1}{G'_1} + \frac{F'_3 - 1}{G'_1 G'_2} + \frac{F'_4 - 1}{G'_1 G'_2 G'_3}$$

which can be written as

$$\begin{aligned} F'_{comp} &= L'_1 + L'_1 (F'_2 - 1) \\ &+ \frac{L'_1 (L'_3 - 1)}{G'_2} + \frac{L'_1 L'_3 (F'_4 - 1)}{G'_2} \end{aligned} \quad (11)$$

The $(S/N)_{out}$ for the system shown in Figure 2 can be shown to be

$$\begin{aligned} \left(\frac{S}{N}\right)'_{out} &= \frac{S'_{in}(1/L'_1)G'_2(1/L'_3)G'_4}{(1/L'_1)G'_2(1/L'_3)G'_4(kT_{sys}W)} \\ &= \frac{S'_{in}}{(kT_{sys}W)} \end{aligned} \quad (12)$$

at the pre-detection output. By using equations (3) and (5), equation (12) can be written as

$$\left(\frac{S}{N}\right)'_{out} = \frac{S'_{in}}{k(RT_0 + T'_{comp})W} \quad (13)$$

For T_{comp} , we substitute equation (11) into (2) and get

$$\begin{aligned} T'_{comp} &= \left[L'_1 F'_2 + \frac{L'_1 (L'_3 - 1)}{G'_2} + \frac{L'_1 L'_3 (F'_4 - 1)}{G'_2} - 1 \right] T_0 \end{aligned} \quad (14)$$

Substituting (14) into (13) yields

$$\left(\frac{S}{N}\right)'_{out} = \frac{S'_{in}}{k\sigma' T_0 W} \quad (15)$$

where σ' is the *noise enhancement* given by

$$\sigma' = R + L'_1 F'_2 + \frac{L'_1 (L'_3 - 1)}{G'_2} + \frac{L'_1 L'_3 (F'_4 - 1)}{G'_2} - 1 \quad (16)$$

Note that if G'_2 is large, then

$$\sigma' \approx R + L'_1 F'_2 - 1 \quad (17)$$

SNR Improvement

The SNR improvement Q is defined as

$$Q \equiv \frac{(S/N)'_{out}}{(S/N)_{out}} \quad (18)$$

Using equations (15) and (9), we derive Q to be

$$Q = \frac{\sigma}{\sigma'} = \frac{(R-1) + L_1 F_2}{(R-1) + L'_1 F'_2 + \frac{L'_1 (L'_3 - 1)}{G'_2} + \frac{L'_1 L'_3 (F'_4 - 1)}{G'_2}} \quad (19)$$

If G'_2 is large, then (19) can be approximated as

$$Q \approx \frac{(R-1) + L_1 F_2}{(R-1) + L'_1 F'_2} \quad (20)$$

Since CDMA has reverse-link power control, the improvement in SNR is manifested in the reduction of mobile transmit power. This is intuitive because if there is an improvement in SNR, reverse power control would command the mobile to power down so that the mobile transmit power is just enough to achieve a desired link SNR. Figure 3 shows the SNR improvement as a function of reverse-link rise and LNA NF. As expected, as reverse-link rise increases (i.e., more noise on the reverse link), the SNR improvement decreases. Also note that the SNR improvement increases as the LNA NF decreases.

Figure 4 shows the SNR improvement as a function of LNA gain and reverse-link rise. As expected, as reverse-link rise increases (i.e., more noise on the reverse link), the SNR improvement decreases. When the LNA gain increases, the SNR improvement also increases; however, the SNR improvement seems to saturate when the LNA gain reaches high levels. Both Figures 3 and 4 show that SNR improvement is more sensitive to changes in rise than changes in amplifier gain.

Capacity Improvement

It is well recognized that the capacity of a CDMA carrier is inversely proportional to the required E_b/N_0 . Therefore, although the link SNR is improved by installing LNAs at the base station, the required E_b/N_0 has not changed. Since the required E_b/N_0 has not changed, the capacity of a base station with LNAs has not improved.

However, the capacity of a neighboring base station (with no LNAs) may improve due to a smaller loading factor. This effect is not easily quantified, however, and the extent of the capacity improvement depends on the network configuration and the distribution of mobiles in the network. Figure 5 presents a possible configuration of LNA deployment. If one assumes that the transmit power of those mobiles in a cell with LNAs is reduced by α , the reverse-link rise is only contributed by those mobiles in the first tier, and mobiles are uniformly distributed, then it can be shown that the average reduction of reverse-link rise is

$$\frac{\Delta \bar{R}}{\bar{R}} = \frac{2(1-\alpha)}{7} \quad (21)$$

Note that if the transmit power is not reduced (i.e., $\alpha=1$), then there is no reduction of the reverse-link rise. On the other hand, the maximum rise reduction is bounded by $2/7$.

Conclusion

The paper has described an analytical model of coverage enhancement as a result of using LNAs. The SNR improvement is found to be more dependent on rise than on the amplifier gain. It also describes ways of deploying LNAs in an existing network to reduce reverse-link rise, thus increase reverse-link capacity. For a more complex network, one can use network simulation to ascertain how much rise can be reduced by deploying LNAs in a certain pattern.

REFERENCES

- [1] Sklar, B. *Digital Communications: Fundamentals and Applications*, Englewood Cliffs N.J.: Prentice-Hall, Inc., 1988.
- [2] Yang, S. *CDMA RF System Engineering*, Norwood MA: Artech House, 1998.
- [3] TIA/EIA IS-95A, "Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband

Spread Spectrum Cellular System,"
Telecommunications Industry Association.

- [4] TIA/EIA IS-98, "Recommended Minimum Performance Standards for Dual-Mode Wideband Spread Spectrum Cellular Mobile Station," Telecommunications Industry Association.

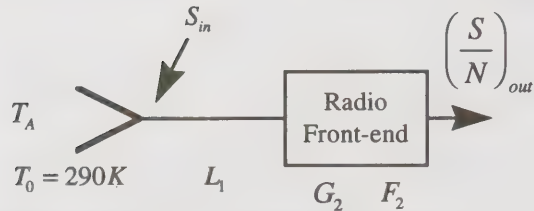


Figure 1. Model of a Base Station Radio Front-End

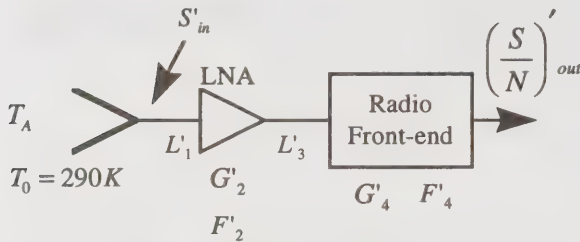


Figure 2. Model of a Base Station Radio Front-End with LNA

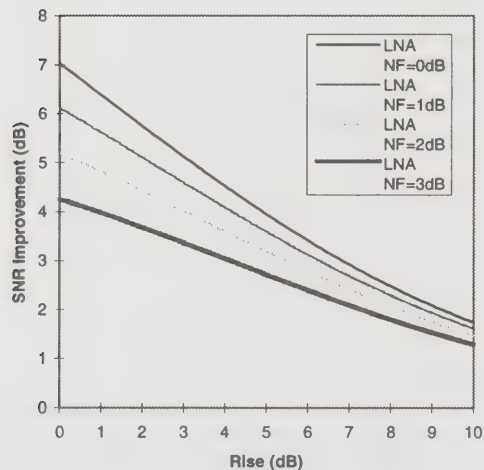


Figure 3. SNR improvement as a function of rise and LNA NF. The graph is generated using the following parameters: $L_1 = 2.0$ dB, $F_2 = F'_4 = 6.0$ dB, $G'_2 = 16$ dB, $L'_1 = 0.5$ dB, and $L'_3 = 1.5$ dB. [2]

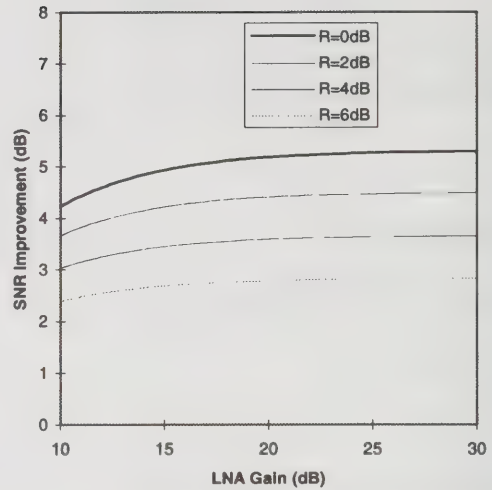
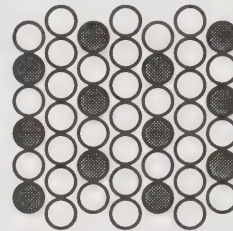


Figure 4. SNR improvement as a function of LNA gain and rise. The graph is generated using the following parameters: $L_1 = 2.0$ dB, $F_2 = F'_4 = 6.0$ dB, $F'_2 = 2.2$ dB, $L'_1 = 0.5$ dB, $L'_3 = 1.5$ dB = 1.41. [2]



- Cell with LNA
- Cell with no LNA

Figure 5. LNA deployment configuration

IS-95 CDMA Forward Link Optimization Tool

John Payne
AirTouch Cellular
3 Park Plaza
Irvine, CA, 92614
(714) 222-8878
John.Payne@AirTouch.com

Asim Qureshi
AirTouch Cellular
3 Park Plaza
Irvine, CA, 92614
(714) 222-8079
Asim.Qureshi@AirTouch.com

Abstract

A processing tool is described that enables the optimization of pilots at a particular traffic load. With the advent of CDMA (Code Division Multiple Access) RF coverage field instruments and a simple technique to characterize a cell's traffic channel power with its traffic load one is able to reconstruct the forward link RF layers on a per pilot basis. By properly combining the RF layers, an accurate picture of the system under load can be obtained which can be used to optimize pilot coverage and diagnosis link or parameter problems. This tool allows the use of field and traffic data for CDMA forward link optimization.

Key words: IS-95 CDMA, Power vs. Erlang Usage, Forward Link Model, Pilot Scanners

I. Introduction

In a typical IS-95 CDMA system, mobiles respond to changes in the RF environment very quickly. The dynamics of the RF environment can be attributed to changes in load, multipath, system parameter settings, and power densities among other things. The current tools available to predict this dynamic condition accurately at the mobile operational level each have their drawbacks. Simulations typically identify coverage holes, i.e. link problems, and load carrying capability of the system. The accuracy of the simulations is dependent upon the precision of the path loss, traffic power usage, and traffic distribution data. Mobile drive data provide a snap shot of information on voice quality, power measurements, and messaging between the mobile and infrastructure. What is not given is the particular load condition that produced the measured interference level.

In this paper we present a tool that combines measured field data of receive power and pilot strengths with measured forward traffic channel power as a function of load at a cell/sector. With this set of information the forward RF link on a per cell/sector basis can be reconstructed. By combining the data for a cluster of cells/sectors a view of the forward coverage

can be obtained with all RF layers accounted for at a particular load level. With this data, pilot coverage can be determined and traded off with other pilots to achieve a desired effect. Also with the data, forward link budget can be calculated to reveal any link problems for the load. The modeled forward link can be compared with drive data to determine if a problem in an area is attributed to link budget or parameter setting deficiencies.

This paper is organized as follows. Section II reviews CDMA forward link equations for the pilot and traffic channels. Section III describes common measurement capabilities of RF pilot scanners. Section IV presents the measurement technique of cell/sector's forward traffic channel power and its correlation to hourly Erlang usage. Section V presents the development of the forward link optimization tool along with an example. Conclusions are given in Section VI.

II. CDMA Forward Link Equations

The forward link for an IS-95 CDMA system is composed of the following code channels: pilot channel, sync channel, paging channel, and traffic channels. The pilot code channel is a binary pseudo-

random (PN) code sequence that is quadrature modulated. All cells transmit the same pilot code sequence but at different chip offset which provides identification of the serving cells for the mobile. The pilot channel facilitates acquisition of the system by the mobile and provides a reference for coherent demodulation. The overhead channels, sync and paging, provide the mobile with system timing and messages.

The primary measurement of a pilot's strength at the mobile is its signal to noise ratio, which is expressed as [1]:

$$\frac{E_c}{I_o} = \frac{P_p * L_p * \frac{1}{R_c}}{P_{Total} * \frac{1}{W_{ss}}} \quad (1)$$

where E_c/I_o is the energy per chip per interference spectral density, P_p is the transmitted pilot power, L_p is the path loss, and R_c is the chip rate of the pilot's PN code. P_{Total} is the total received power and W_{ss} is the bandwidth of the CDMA channel. In IS-95, R_c is 1.2288 Mcps and W_{ss} is 1.25 MHz, which reduces (1) in decibels (dB) to:

$$\frac{E_c}{I_o} (dB) = P_p (dB) + L_p (dB) - P_{Total} (dB) \quad (2)$$

Equation (2) specifies the pilot receive power at the mobile is a relative measurement to the total receive power in dB [2].

For the forward link traffic channels, the figure of optimization is the energy per bit per interference spectral density, E_b/I_o . E_b/I_o is calculated as:

$$\frac{E_b}{I_o} = \frac{P_{Tch} * L_p * \frac{1}{R_b}}{P_{Total} * \frac{1}{W_{ss}}} \quad (3)$$

where P_{Tch} is the transmitted traffic channel power and R_b is the data bit rate. The ratio of W_{ss}/R_b is defined as the processing gain (PG). Converting units to dB, equation (3) becomes:

$$\frac{E_b}{I_o} (dB) = P_{Tch} (dB) + L_p (dB) - P_{Total} (dB) + PG \quad (4)$$

III. RF Pilot Scanners

Within the past year and a half, test instrument manufacturers have developed CDMA RF coverage instruments, i.e. pilot scanners [3], that accurately measure pilot strengths (E_c/I_o), pilot energy (E_c), receive power, and chip delay information all within a 30-msec interval. This data can be tagged with geographical information to obtain coverage layout [3]. Power measurement accuracy is typically within ± 2 dB and chip delay accuracy is within 1 chip duration. Measurement of E_c/I_o values relative to total received

power can be made within a specified window, typically within 18 dB.

With this set of information, pilot receive power and pilot path loss information can be derived by using equation (2) as follows:

$$P_{p_Rx} (dB) = \frac{E_c}{I_o} (dB) + P_{Total} (dB) \quad (5)$$

where P_{p_Rx} is the pilot received power at the mobile and is defined as:

$$P_{p_Rx} (dB) = P_p (dB) + L_p (dB) \quad (6)$$

The path loss, L_p , can be derived using equation (6).

IV. Forward Traffic Channel Power vs. Erlang

Power measurements were taken at the output of a cell's linear power amplifiers (LPA) by connecting directional couplers in the RF path to the antennas. The power meter used has the capability of scanning through the CDMA spectrum in less than 20 seconds thereby taking measurements on three sectors within a minute. Figure 1 shows a typical 24-hour measurement cycle at a cell site. These particular measurements were taken in 1.5-minute intervals. The rise in power is attributed to traffic channels being used while the base of the graph represents the cell baseline power, which is composed of power for the pilot, paging and synch channels.

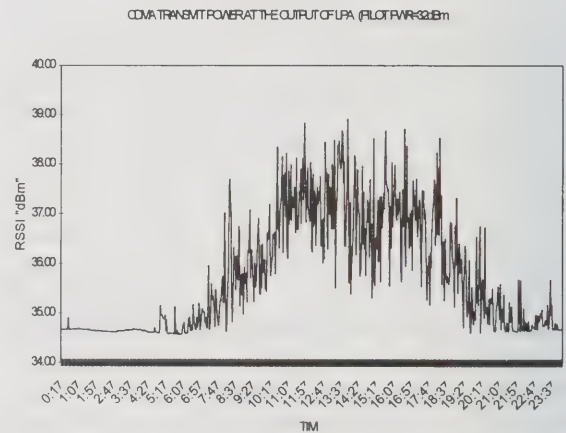


Figure 1 Change in traffic channel power over a 24-hour period.

For each hourly Erlang interval, the average, maximum, and minimum traffic channel power data points were calculated and graphed against the hourly usage. The data points were typically collected over several weeks. This measured traffic channel power when correlated with per hour Erlang usage shows a

linear relationship as seen in Figures 2. Figure 2 shows an example measurement for a site where the maximum power level has 0.32-dB/Erlang slope and the average power level has a slope of 0.22-dB/Erlang. A dozen sites were measured using the same technique and the linear relationship was found to hold true. The only measurement difference between the cells was a change in slope and intercept, which was expected.

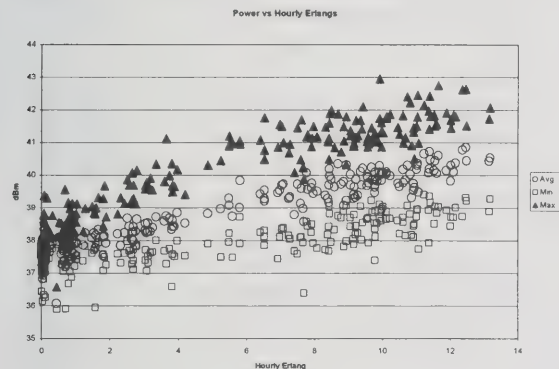


Figure 2 Traffic Channel Power vs Hourly Erlang Usage

This method of correlating traffic channel power to hourly Erlang usage facilitates the modeling of traffic load with a particular traffic channel power with reasonable accuracy.

V. Forward Link Optimization Tool

Since E_c/I_o values must be measured within a defined window relative to total received power, care must be taken that the total received power for the area of concern is minimal to obtain accurate E_c/I_o values to derive path loss information with in acceptable limits. For example, for a E_c/I_o window of 18.5 dB it is recommended that the total received power level be in the range of -80 to -85 dBm in order to obtain minimum pilot receive power in the -100 to -105 dBm levels. This may require data collection during non-busy time and/or performing iterative data collection of desired pilots while powering down other cells and repeating the process for other pilots in the desired area.

Figure 3 depicts an example of an area between 3 cells where modeling is desired. Pilot data and channel receive power is ideally collected from all twelve cells when no traffic channel power are present. The following steps outline the initial data processing needed to compute pilot path loss information:

- Step A:** Divide desired area into bins.
- Step B:** For each bin calculate average receive power for each pilot detected within the bin. Compute path loss for each detected pilot using equation 5 and 6, and pilot effective radiated power (ERP).

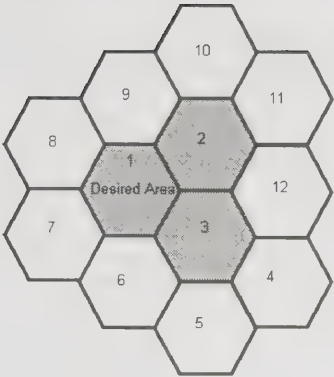


Figure 3 Modeled area of 3 cells.

Table 1 shows an example of a set of bins and the results obtained from following the above steps to derive path loss data.

Bin1: Pilot A ERP = 44 dBm Pilot B ERP = 30 dBm Measured Parameters: I_o = -78 dBm Pilot A: E_c/I_o = -8 dB Pilot B: E_c/I_o = -10 Calculated Path Loss: Path Loss of Pilot A = -130 dB Path Loss of Pilot B = -118 dB	Bin2: Pilot A ERP = 44 dBm Pilot B ERP = 30 dBm Pilot C ERP = 41 dBm Measured Parameters: I_o = -85 dBm Pilot A: E_c/I_o = -10 dB Pilot B: E_c/I_o = -12 dB Pilot C: E_c/I_o = -11 dB Calculated Path Loss: Path Loss of Pilot A = -139 dB Path Loss of Pilot B = -127 dB Path Loss of Pilot C = -137 dB
--	---

Table 1 A 2-bin example of calculating path loss information.

Once completed, modeling the area under traffic load requires a simple accounting of all the forward link powers from each cell. To achieve that aim, the following information must be gathered for each cell: pilot effective radiated power (ERP), overhead power (ERP), and traffic load power (ERP) which is derived from measured power from the cell as correlated to Erlang usage as shown in section IV. The following steps outline the procedure to model the forward link under a traffic load condition at each bin:

- Step 1:** For each pilot within a bin calculate pilot received power as shown in equation 6.
- Step 2:** Calculate total associated forward link power for each pilot, i.e. include pilot overhead power and traffic channel power.
- Step 3:** Add all powers linearly. This value becomes the new I_o or total receive power for the bin.

Step 4: Compute new E_c/I_o values for each pilot using calculated pilot receive power and the new total receive power value obtain in Step 3 using equation 2.

Table 2 shows the forward link under load results using the above steps for the example in Table1.

Bin1: Pilot A ERP = 44 dBm Pilot A Overhead = 3 dB Pilot A Traffic Power = 4 dB Pilot B ERP = 30 dBm Pilot B Overhead = 3 dB Pilot B Traffic Power = 5 dB Total Power per Pilot: Cell A Total Power = -79 dBm Cell B Total Power = -80 dBm Total Rx Pwr (I_o)=-76.5 dBm New E_c/I_o: Pilot A E_c/I_o = -9.5 dB Pilot B E_c/I_o = -11.5 dB	Bin2: Pilot A ERP = 44 dBm Pilot B ERP = 30 dBm Pilot C ERP = 41 dBm Pilot A Overhead = 3 dB Pilot B Overhead = 3 dB Pilot C Overhead = 3 dB Pilot A Traffic Power = 4 dB Pilot B Traffic Power = 5 dB Pilot C Traffic Power = 3 dB Total Power per Pilot: Cell A Total Power = -88 dBm Cell B Total Power = -89 dBm Cell C Total Power = -90 dBm Total Rx Pwr(I_o)=-84.15 dBm New E_c/I_o: Pilot A E_c/I_o = -10.85 dB Pilot B E_c/I_o = -12.85 dB Pilot C E_c/I_o = -11.85 dB
--	---

Table 2 Forward link model for 2-bin example.

Sometimes it is not possible to collect data for all the cells to complete the model e.g. cells 4-12 in figure 3, therefore a baseline noise floor must be determined. For the time period in question receive power data (I_o) must be collected for the desired area. For each bin, calculate the difference between calculated total receive power from the I_o collected. This difference becomes the noise floor for the bin. Table 3 demonstrates the calculations and changes to pilot strengths.

Bin 1: Calc.TotalRxPwr(I_o)=-76.5 dBm I_o Measured = -74 dBm Calc. Noise Floor=-77.6dBm New E_c/I_o: Pilot A E_c/I_o = -12 dB Pilot B E_c/I_o = -14 dB	Bin 2: Calc.TotalRxPwr(I_o)=-84.15 dBm I_o Measured = -82 dBm Calc. Noise Floor=-86.1dBm Pilot A E_c/I_o = -13 dB Pilot B E_c/I_o = -15 dB Pilot C E_c/I_o = -14 dB
--	---

Table 3 A 2-bin example of noise floor calculation.

After completing all necessary bin calculations, pilot powers can be adjusted to see a shift in coverage. This process requires recalculation of the total power per pilot and total receive power (I_o) per bin. The changes in coverage will affect hourly Erlang usage for affected cells, which must be estimated [2]. These Erlang values will drive traffic load power to use in the model. It is recommended using the same traffic channel power per Erlang slope at high power levels with a correction factor of +/- 2dB.

For each bin, link budget calculations can be performed to determine if the forward link can be closed. By using equation (4) and assuming a target E_b/N_o , e.g. 10-dB, along with the total receive power and path loss data per bin, the transmitted traffic channel power, P_{Tch} , can be calculated and compared to the maximum and minimum traffic channel power requirements for the serving cells. If P_{Tch} is within the power requirements, the forward link can be closed. Comparison of this data per bin with drive data of problems spots can determine if the cause is link budget related or a parameter setting problem.

VI. Conclusion

The forward link optimization tool described has many advantages such as optimizing coverage and soft-handoff, identifying system parameter problems, and possibly predicting coverage holes under load. But the tool has disadvantages due to its need for large sum of data to complete an accurate model. Further investigation is needed to understand the trade-off between minimizing data collection and model precision. Additional research is needed to fully understand the mechanics behind the linear relationship between traffic channel power versus usage and how power control, soft-handoff, and interference levels influence the traffic channel power/Erlang slope.

Reference:

- [1] TIA/EIA/IS-98 "Recommended Minimum Performance Standards for Dual-Mode Wideband Spread Spectrum Cellular Mobile Stations"
- [2] "Applications of CDMA in Wireless/Personal Communications," by V. K. Garg, K. Smolik, J. Wilkes, 1997 Prentice Hall
- [3] "HP E7450B, HP E7452A, HP E7460A, HP E7470A, and HP E7472A Drive Test Solutions Quick Start Guide" 1997 Hewlett-Packard.
- [4] "CDMA Principles of Spread Spectrum Communication," by A. J. Viterbi, 1995 Addison-Wesley.

Simulation of Coded MC-DS-CDMA Systems

Thad B. Welch[†] and Rodger E. Ziemer[‡]

[†]United States Naval Academy

Department of Electrical Engineering

Annapolis, MD 21402-2005

(410) 293-6163 or (410) 293-3493 (fax) or t.b.welch@ieee.org

[‡]University of Colorado at Colorado Springs

College of Engineering & Applied Science

Electrical and Computer Engineering Department

P.O. Box 7150

Colorado Springs, CO 80933-7150

(719) 262-3350 or (719) 262-3589 (fax) or ziemer@eas.uccs.edu

Abstract

We use computer simulations to compare the performance improvements associated with coding on a gaussianized BPSK-MC-DS-CDMA system and a non-gaussianized DPSK-MC-DS-CDMA system. The effects of coding on the gaussianized system approach the performance of a coded BPSK system operating in a Gaussian channel while the non-gaussianized system performance is approximately 4 dB from the coded DPSK system operating in a Gaussian channel.

Introduction

Several authors, for example [1–5], propose and analyze multicarrier direct-sequence code-division multiple-access (MC-DS-CDMA) systems for use in a mobile radio environment. In these systems, if the bandwidth of a wideband direct-sequence spread-spectrum (DSSS) system exceeds the coherence bandwidth, $(\Delta f)_c$, of the channel, the signal is severely distorted [6]. This type of channel is said to be frequency-selective and a RAKE receiver [7] can be used to enhance system performance [8]. If a single baseband DSSS signal is modulated by multiple subcarriers that are properly separated, these subcarriers will experience independent fading. With multiple replicas of the message signal being transmitted through independent fading

channels, diversity gain can be achieved by the receiver. A maximal ratio combiner (MRC) can be used to recombine these message signals to improve the system's error performance [9]. Kondo and Milstein [4], propose and analyze such a system and show that their MC-DS-CDMA system performance with M subcarriers is identical to a single carrier RAKE receiver system occupying the same bandwidth and processing M resolvable paths ($L_1 = M$). As M increases, the bit error probability (BEP) associated with this system asymptotically approaches the performance for a binary phase shift keying (BPSK) system operating in an additive white Gaussian noise (AWGN) channel. This *gaussianization* of the system's performance is shown in Fig. 1.

A more easily realized system (DPSK-MC-DS-CDMA) was proposed and analyzed in [5]. The receiver and transmitter block diagrams are shown in Fig. 2 and Fig. 3. In Fig. 2,

Work of R.E. Ziemer supported in part by the Office of Naval Research under contract N0014-920-J01761/P00004

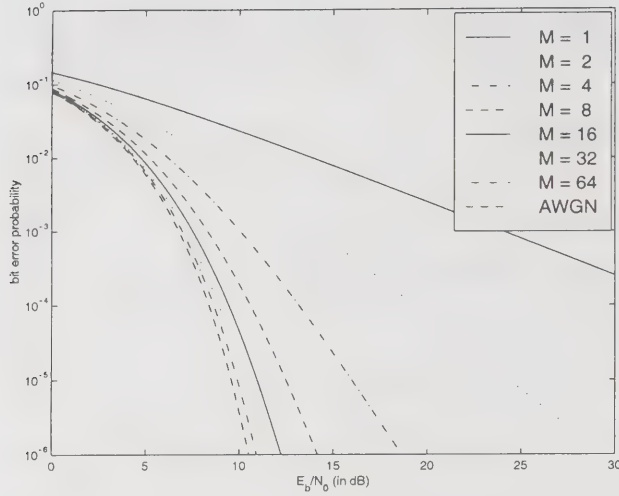


Fig. 1. BEP for Kondo and Milstein's MC-DS-CDMA system.

differential phase shift keying (DPSK) is used to help reduce the effect of the Doppler-induced rapid phase variations of the received signal. The DSSS spreading process is shown as the multiplication of the DPSK data with the local pseudo-noise (PN) code. The multicarrier local oscillators are at frequencies f_1, f_2, \dots, f_M , where M is the total number of subcarriers. The wireless channel block represents the effect of Doppler spreading. Finally, the I&Q AWGN are the inphase and quadrature components of an additive white Gaussian noise process at the front end of the receiver. In Fig. 3, the DSSS despreading process is shown as the multiplication of the received signal with the local estimate of the PN code. I&D represents the receiver integrate-and-dump function which is synchronized to the bit transition times. The delay* blocks represent the delay and multiply associated with the DPSK decode process and the * is used to denote complex conjugation.

This system is more easily realized because it is differentially coherent and utilizes an equal gain combiner (EGC). The BEP curves associated with this system are shown in Fig. 4. As the number of subcarriers is increased, the system's BEP does not closely approach the BEP associated with an AWGN channel; however, the shape is very similar to a Gaussian per-

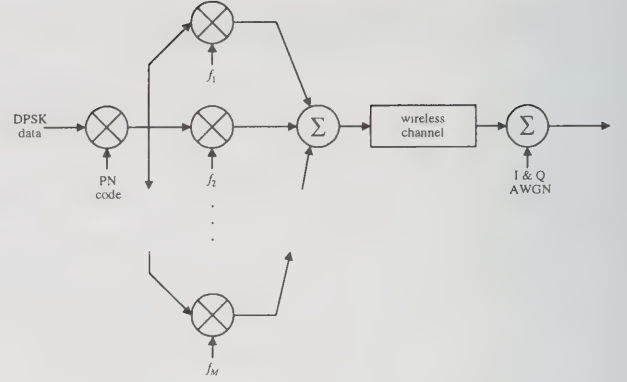


Fig. 2. Transmitter block diagram from [5].

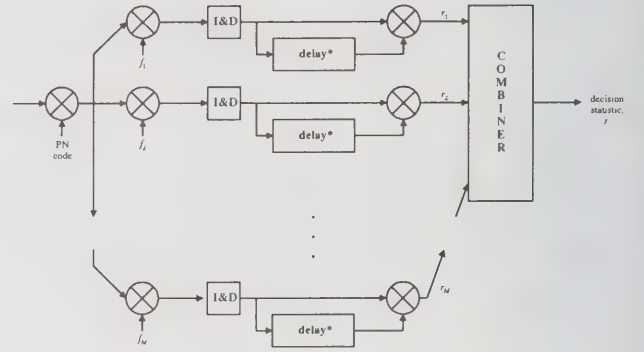


Fig. 3. Receiver block diagram from [5].

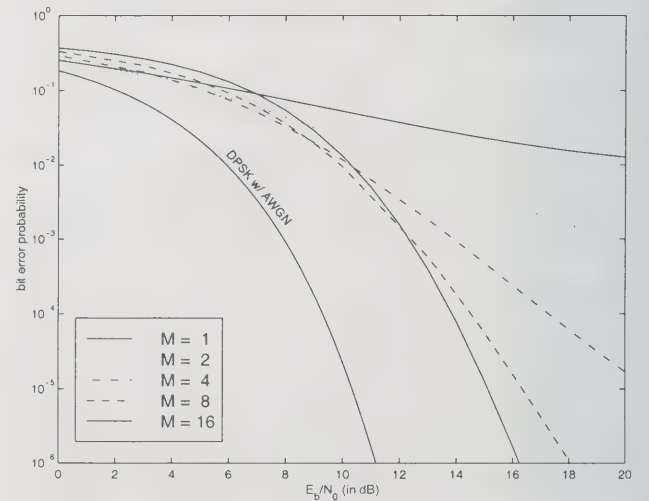


Fig. 4. BEP curves associated with a differentially coherent MC-DS-CDMA system with EGC.

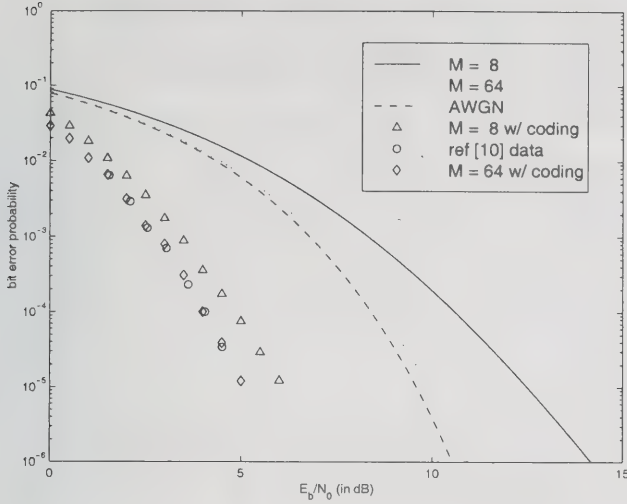


Fig. 5. BEP for an uncoded and coded systems.

formance curve. This system exchanges BEP performance for ease of implementation.

The Addition of Coding

In [4], as M increases and the fading channel BEP approaches the BEP associated with an AWGN channel, it is reasonable to expect that the introduction of convolutional coding with soft Viterbi decoding will also approach AWGN coded channel performance. This assumption was verified, and is shown in Fig. 5, where the results for $M = 64$, using a convolutional code with rate, $r = \frac{1}{2}$, and constraint length, $K = 3$ are shown as diamonds. The difference between the results in [10] (results shown as circles) and our simulation¹ results are less than 0.1 dB. Also note that the simulation results for the coded $M = 8$ case exhibits significant coding gain. Less than 1 dB separates the $M = 8$ and the $M = 64$ coded results.

For the system analyzed in [5], a Doppler spread channel is assumed. Since a Doppler spread channel exhibits memory, the introduction of convolutional coding should also be accompanied by an interleaver. An interleaver will *breakup* the burst errors associated with a deep channel fade and allow the convolutional code to be more effective. Since a infinite

¹All simulations were conducted using MATLAB version 5.2 with the communications and signal processing toolboxes.

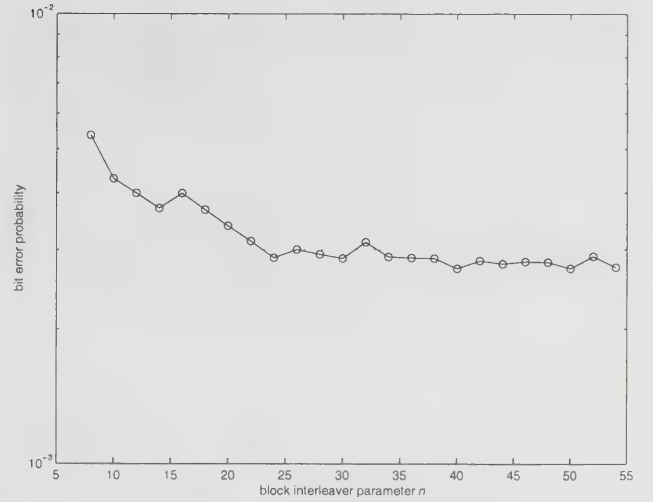


Fig. 6. Bit error probability as a function of the interleaver parameter n .

depth interleaver is neither practical or capable of being simulated, a compromise in interleaver depth is needed. A block interleaver of depth $n \times n$ was selected with the value of n determined by simulation. A repeatable² simulation was conducted, varying only the parameter n . The results of these simulations³ are shown in Fig. 6. As shown in Fig. 6, for $n \geq 24$ the improvement in system error performance is small. Additionally, for $n > 10$ the throughput of the computer simulation as measured in simulated bits per second decreases rapidly. For these reasons a 24×24 block interleaver was selected.

As shown in Fig. 4, $M = 8$ is a reasonable choice for the number of subcarriers for operation in the $10^{-2} \rightarrow 10^{-3}$ region. The selection of $M = 8$ will also allow for a comparison to the coded performance of [4] as shown in Fig. 5. The results of these simulations are shown in Fig. 7. The simulation results for the $M = 8, r = \frac{1}{2}, K = 3$, and $f_d T = 0.04$ asymptotically approaches the BEP expected for an AWGN channel. The bounds for the BEP equations for convolutionally coded BPSK with soft Viterbi decoding are discussed in [11]. We

²The *seed* associated with the random number generators was always initialized at the same value.

³Throughout the computer simulations, the product of the maximum Doppler frequency, f_d and the bit period, T remained constant at $f_d T = 0.04$.

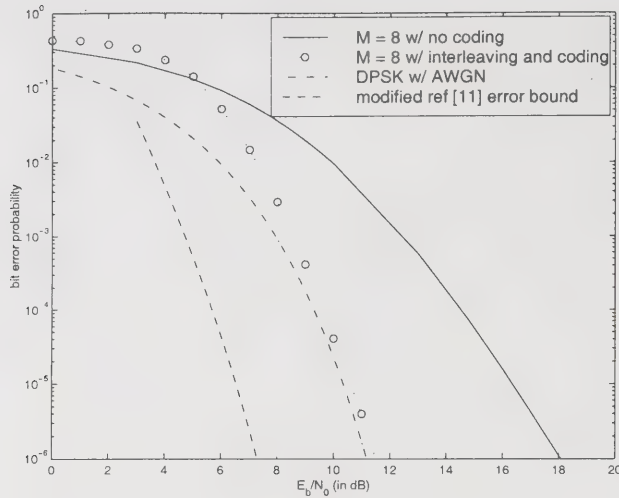


Fig. 7. BEP for uncoded and coded systems.

modified the equation to allow for coded DPSK by changing

$$P_E < \sum_{k=d_{free}}^{\infty} c_k Q\left(\sqrt{\frac{2k r E_b}{N_0}}\right)$$

to

$$P_E < \sum_{k=d_{free}}^{\infty} c_k \exp\left(-\frac{k r E_b}{N_0}\right)$$

where d_{free} is the free distance, c_k is a partial weight coefficient, and $Q(\cdot)$ is the Q-function of (\cdot) . Tables listing the appropriate values for d_{free} and c_k for $r = \frac{1}{2}$ and $K = 3 \rightarrow 9$ can be found in [11]. The actual coding gains achieved and the expected coding gains for an AWGN channel are listed below,

BEP	Actual gain	AWGN gain
10^{-2}	2.6 dB	2.3 dB
10^{-3}	3.8 dB	3.2 dB
10^{-4}	4.9 dB	3.6 dB
10^{-5}	5.7 dB	3.8 dB

Conclusions

In this paper we have used computer simulations to compare the performance improvements associated with convolutional coding with soft Viterbi decoding on a *gaussianized* MC-DS-CDMA system [4] and a more easily realized DPSK-MC-DS-CDMA system which was initially analyzed in [5]. Specifically, we have

shown that the *gaussianized* MC-DS-CDMA system will approach AWGN coded channel performance. While the DPSK-MC-DS-CDMA system achieves more coding gain than a single carrier DPSK system operating in an AWGN channel, its performance is offset by approximately 4 dB from a coded DPSK operating in an AWGN channel. The system's performance also asymptotically approaches the performance expected of an *uncoded* single carrier DPSK system operating in an AWGN channel. Finally, for smaller values of M , where the system exhibits an irreducible BEP, adding coding can significantly reduce or even eliminate this error floor.

REFERENCES

- [1] E. Sourour and M. Nakagawa, "Performance of orthogonal multi-carrier CDMA in a multipath fading channel," *IEEE Trans. on Commun.*, vol. 44, pp. 356-367, Mar. 1996.
- [2] M. Ohkama, R. Kohno, and H. Imai, "Orthogonal multicarrier FH-CDMA schemes for frequency selective fading," *IEICE Trans. on Commun.*, vol. J77-B11, pp. 691-702, Nov. 1994.
- [3] R.N. Yee and J.P. Linnartz, "Controlled equalization of multi-carrier CDMA in an indoor Rician fading channel," *Proc. VTC'94*, pp. 1665-1669, 1994.
- [4] S. Kondo and L.B. Milstein, "Performance of multi-carrier DS CDMA systems," *IEEE Trans. on Commun.*, vol. 44, pp. 238-246, Feb. 1996.
- [5] T.B. Welch and R.E. Ziemer, "DSSS/MCM system performance in a Doppler spread channel," *Proc. MILCOM'97*, pp. 1-5 of section 16-05 [CD-ROM], Nov. 1997.
- [6] J.G. Proakis, *Digital Communications*, McGraw-Hill, Inc., New York, 1995.
- [7] R. Price and P.E. Green, "A communication technique for multipath channels," *Proc. IRE*, vol. 46, pp. 555-570, Mar. 1958.
- [8] G.L. Turin, "Introduction to spread-spectrum anti-multipath techniques and their application to urban digital radio," *Proc. IEEE*, vol. 68, Mar. 1980.
- [9] M. Schwartz, W.R. Bennett, and S. Stein, *Communication Systems and Techniques*, IEEE Press, New Jersey, 1996.
- [10] J.A. Heller and I.M. Jacobs, "Viterbi decoding for satellite and space communications," *IEEE Trans. on Commun.*, vol. 19, pp. 84, 1971.
- [11] R.E. Ziemer and R.L. Peterson, *Introduction to Digital Communications*, Macmillan, New York, 1992.

Analysis of Impact on Handset Transmitter Design of the High-Speed Data Requirements in the IS-95-B CDMA Wireless Standard

Giridhar Mandyam
Texas Instruments Inc.
P.O. Box 660199, M/S 8723
Dallas, TX 75266-0199

Ph: (972) 480-6511, Fax: (972) 480-6552, Email: gdm6@msg.ti.com

Abstract

As part of the recent evolution of the IS-95 CDMA wireless communications towards third-generation technology, high-speed data requirements were defined for both the uplink and downlink in the recently proposed IS-95-B revision. Although the requirements for the downlink are straightforward, the uplink requirements have considerably larger impact on the handset transmitter and base station receiver design. In this paper, the impacts on the transmitter are examined, from baseband to RF frequencies. In particular, waveform quality and emissions are examined through high-level simulation to determine the tradeoffs involved in implementation of a handset design which can support IS-95-B high speed data requirements for the uplink in the 1.9 GHz band.

Introduction

The IS-95 standard poses many challenges for implementation, particularly in the design of a proper transmitter. The required design must be linear over a large dynamic range (nearly 100 dB) yet still be designed efficiently so as to compete with other existing wireless standards. Therefore, careful system analysis is oftentimes necessary prior to implementing enhancements to the existing IS-95 functionality.

Recently, the IS-95 standard has been expanded to include high-speed data functionality in both the uplink and downlink [2]. The implementation for multiple user data channels, i.e. traffic channels, is straightforward in the downlink. In this case, the mobile user is already required to demodulate multiple data channels, each defined by a unique Walsh code sequence, in order to establish and maintain a communications link. The advantage of Walsh code sequences is that due to their mutual orthogonality, one particular channel may be extracted while suppressing all other channels at the same time simply by modulating the incoming data with the desired Walsh sequence. Moreover, due to the processing gain of approximately 21 dB in addition to orthogonal Walsh modulation, multiple data channels may be de-

modulated simultaneously without increasing the dynamic range of the receiver.

However, channelization on the uplink is accomplished by time-offsets on the spreading sequences used for digital modulation. Although the baseline transmit modulation scheme involves offset-QPSK, due to the requirement for baseband pulse shaping, the amplitude information of the transmitted data becomes necessary. As a result, transmitting multiple data channels on the uplink results in an increase in the dynamic range which the transmitted data occupies. This can result in increased complexity in the analog baseband and RF circuitry.

The purpose of this work is twofold: (1) This work is intended to provide a high-level analysis of the IS-95 handset transmitter under high-speed data applications, and (2) This work is intended to specify the tradeoffs between complexity of analog IC design and the number of high-speed data channels supported in an IS-95 compliant handset.

High Speed Data Requirements

In order to understand high speed data requirements, the physical layer requirements of IS-95-A should be examined. In this standard, only

one channel transmission is supported on the uplink. The basic physical layer implementation of the mobile transmitter digital baseband section is pictured in Figure 1. As shown, the output of the convolutional coder (rate 1/2 or rate 1/3) undergoes interleaving, symbol repetition, and puncturing (not shown) and is passed to a 64-ary Walsh modulator. This is simply a mapping of every 6 symbols to a 64-bit Walsh code sequence. This type of modulation aids in noncoherent detection at the base station. The sequence is then spread by a 1.2288 Msps pseudonoise sequence known as the *long code*. This sequence is generated from a length-42 linear feedback shift register (LFSR), and thus has a period which is long in duration. Then, the sequence undergoes quadrature spreading with 1.2288 Msps pseudonoise sequences on both the I and Q channels (denoted PN_I and PN_Q , respectively). Finally, after shifting the Q channel to accomplish offset-QPSK modulation, both channels are passed through a 48-tap FIR whose tap spacing is a quarter-chip ("chip" refers to the duration of a bit of the pseudonoise sequence). Due to this type of baseband filtering, the amplitude information has become important, and thus the transmitter behaves linearly over the dynamic range of the signals output from the filters.

The output of these filters feeds into digital-to-analog converters (DAC's) and into other analog baseband and RF circuitry. A top-level diagram of this part of the mobile transmitter is shown in Figure 2.

The addition of multiple data channels in the uplink would require these channels to simply be added up prior to digital-to-analog conversion. Although the channelization is accomplished by different time offsets on the long code, IS-95-B supports the use of multiple phases on the I and Q data so as to reduce the peak-to-average ratio. In other words, given the QPSK mapping of the fundamental data channel (0,1 to 1,-1), all additional data channels are mapped with a relative phase offset to the fundamental data channel constellation. These phases are listed in Table I.

Although the digital baseband implementation of a handset transmitter must change to accommodate high speed data, it is desirable to limit the change in analog baseband and RF circuitry to support this enhancement.

Modeling and Simulations

The modeling depicted in Figure 1 and Figure 2 was performed so as to properly simulate a CDMA system using MATLAB. Random bits were generated and passed through the baseband filters, which were modeled with the FIR coefficients quantized to 9 bits. The 8-bit DAC's on both the I and Q channels were modeled with .9 bits of differential nonlinearity (DNL) and 2 bits of integral nonlinearity (INL). It was also assumed that 32 bits at the output of each DAC were reserved for possible DC-offset correction, and as a result only 224 levels were available at each DAC. Each DAC's nominal output ranged from -.75 volts to .75 volts. These DAC's were followed by image-reject filters, modeled as third-order Butterworth type.

It is important to take note of the scaling of the FIR filter output samples, as this is how one maximizes the use of the available headroom on the DAC. Moreover, one must choose scaling factors that are sufficiently robust regardless of the number of data channels being transmitted; this requires some experimentation.

The RF section modeling was performed starting with the IQ-modulator. This was modeled assuming a certain amount of gain and phase imbalance, in addition to DC offset resulting from baseband. The modulation was accomplished using sinusoids rather than square waves; this was done to avoid having to model a bandpass filter following the modulator. The transmit automatic gain control (AGC) amplifier was modeled using a classic third-order polynomial description employing the input third-order intercept point. The upconverter was modeled as an image reject filter with gain and phase imbalance in the phase shifting networks. Finally, the power amplifier was modeled using the AM-to-AM and AM-to-PM response for a single tone, as discussed in [3]. In this paper, the authors showed the effectiveness of such a model for a PCS PA. The PA characteristics employed for this analysis are shown in Figure 3.

The choice of IF frequency was 120 MHz. This is a good frequency choice for upconversion to the PCS band. A simulated IF spectrum is pictured in Figure 4. Note the spectral images appearing at 5 MHz intervals from the desired signal; these images result from the DAC operation. These im-

ages can be thought of as spurious products which, when given as input to the PA, could result in severe spectral growth if the PA is operating near compression.

As a first test, it was assumed that there was .5 dB of gain mismatch and 3 degrees of phase mismatch collectively in the IQ-modulator and the image-reject filters. Moreover, 3 percent bandwidth mismatch was assumed in the image reject filters, given a nominal corner frequency of 630 kHz. 3 mV of DC offset was assumed to be resulting from baseband. The transmit AGC amplifier was assumed to be running at 10 dB gain and -15 dBm input intercept assuming 500 ω input impedance; this is very close to a worst-case for the input intercept point for a typical CDMA transmit AGC amplifier at this gain setting. Finally, .1 dB of gain mismatch and 1 degree of phase mismatch were assumed in the upconverter, modeled as an image-reject mixer with 50 ω input impedance. In addition, it was assumed that 10 mV of DC-offset was present in the image-reject mixer.

The results for high-speed data simulations (up to 4 channels) are shown in Table II for emissions (measured over a 30 kHz bandwidth at the 1.25, 2.75, and a 1 MHz bandwidth for 4.24 MHz) and ρ , the waveform quality factor. According to the requirements of J-STD-018 [1], the minimum performance specification for mobile handsets supporting the IS-95 standard at PCS frequencies, the minimum waveform quality specified is .944, and the emissions level should be at -42 dBc/30 kHz at 1.25 MHz away from center frequency, and -60 dBm/30 kHz at 2.75 MHz away from center frequency. At 4-channel transmission, the waveform quality and emissions at 1.25 MHz start to approach barely acceptable levels. However, the emissions requirements are met in all 4 cases, given a maximum PA output power of 27 dBm.

The best way to remedy the degradation due to 4-channel transmission without affecting the RF circuitry is to increase the baseband DAC's resolution to 9 bits. An 8-channel simulation was run assuming .9 bits of DNL and 4 bits of INL in the DAC's. In this case, the emissions at were -52.2 dBc, -76.4 dBc, and -75.1 dBc at 1.25, 2.75 and 4.24 MHz, respectively. The waveform quality was .972.

Conclusions

Based on preliminary systems-level simulation results, it can be concluded that there is a considerable cost associated with implementation of IS-95-B high speed data requirements for the uplink if one chooses to support the maximum bitrate provided by 8 data channels. The degradation in performance due to the increase in dynamic range of the data may be offset if one increases the resolution of the I and Q DAC's. Given that DAC'S required are high-frequency (on the order of 5 Msps), this requirement can result in an expensive design revision for existing IS-95 mobile handset architectures.

As IS-95-B is right now only just being approved as a standard, much more insight into the price of supporting high speed data capability will come about during the next couple of years. However, regardless of whether mobile manufacturers will produce handsets which support this capability, much of the knowledge gained by understanding the impact on handset design will be useful in the design of third-generation handsets.

REFERENCES

- [1] *J-STD-018: Recommended Minimum Performance Requirements for 1.8 to 2.0 GHz Code Division Multiple Access (CDMA) Personal Stations*. The Telecommunications Industry Association, Arlington, VA, 1995.
- [2] *TIA/EIA/IS-95-B Ballot Version: Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System*. The Telecommunications Industry Association, Arlington, VA, 1998.
- [3] W. Struble, F. McGrath, K. Harrington, and P. Nagle. Understanding linearity in wireless communication amplifiers. *IEEE Journal of Solid State Circuits*, 32(9):1310-1318, Sept. 1997.
- [4] E. Tiedemann, T. Chen, and O. Glauser. Improving the reverse link peak to average ratio. *TIA TR45.5 Submission TR45.5.3.1/97.06.16.05*, June 1997.

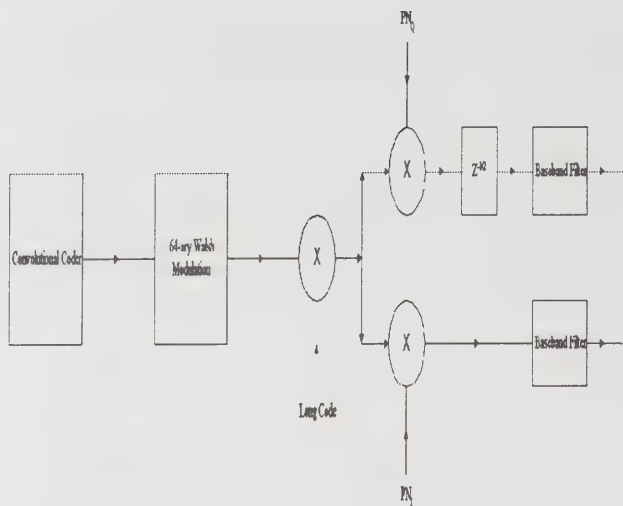


Fig. 1. Uplink Chain

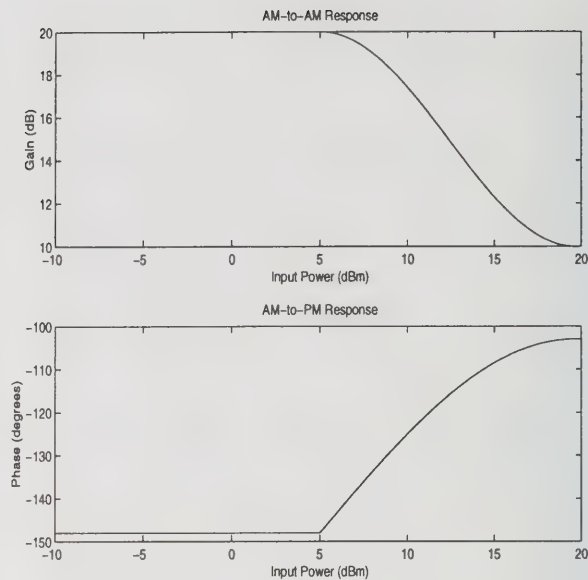


Fig. 3. Power Amplifier Characteristics

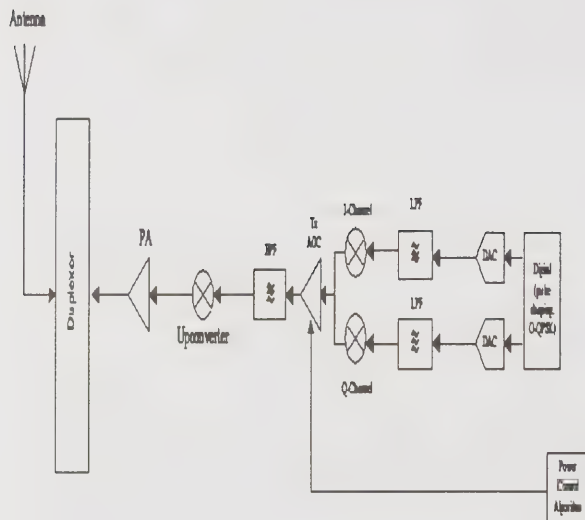


Fig. 2. Mobile Transmitter

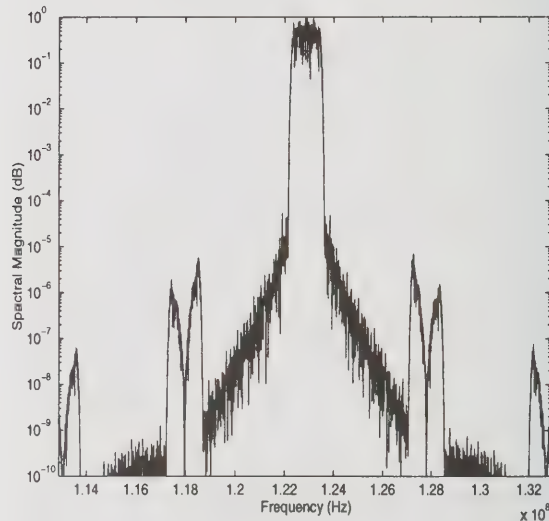


Fig. 4. IF Spectrum

Additional Channel	Phase Offset from Fundamental Channel
1	$\pi/2$
2	$\pi/4$
3	$3\pi/4$
4	0
5	$\pi/2$
6	$\pi/4$
7	$3\pi/4$

TABLE I

High-Speed Data Phasing for IS-95-B

Channels	1.25 MHz	2.75 MHz	4.24 MHz	ρ
1	-55.2 dBc	-86.3 dBc	-74.7 dBc	.984
2	-52.4 dBc	-83.5 dBc	-75.3 dBc	.980
3	-48.9 dBc	-74.1 dBc	-75.6 dBc	.968
4	-46.1 dBc	-72.2 dBc	-74.9 dBc	.953

TABLE II

High-Speed Data Simulation Results

Effects of Carrier Tracking in RAKE Reception of Wideband DSSS

R. E. Ziemer

Department ECE

University of Colorado at Colorado Springs
Colorado Springs, CO 80933, (719)262-3350
(719)262-3589 (fax), ziemer@eas.uccs.edu

B. R. Vojcic

Department of EECS

The George Washington University
Washington, D.C. 20052

L. B. Milstein

Department of ECE

University of California at San Diego
La Jolla, CA 92093

J. G. Proakis

Department of ECE

Northeastern University
Boston, MA 02115

Abstract

An analysis is given of the effect of phase tracking error on the fingers of a coherent RAKE receiver. Results for constant and exponentially decaying power delay profiles indicate that for sufficiently low tracking loop signal-to-noise ratio, a point is eventually reached where performance does not improve with the use of increased number of fingers. This suggests that the expected performance improvement due to finer multipath resolution in the very wideband personal communications systems anticipated for third generation will be negated by parameter estimation error unless special care is taken to minimize these errors, such as optimized parameter estimation procedures.

Introduction

Third generation (3G) personal communication systems (PCS) will invariably be wideband due to the demands of mixed traffic (voice, video, and data), the need for variable rate transmission, and the desire to more finely resolve the channel, thereby achieving more diversity gain. Currently, one of the leading candidate wideband modulation methods is direct-sequence spread-spectrum (DSSS). In regard to achieving more diversity gain, the use of RAKE detection is a leading candidate, although other possibilities, such as space diversity and coding, will also be used in concert with RAKE.

This paper focuses on RAKE reception in a wideband context for improving performance in 3G PCS. Wider bandwidth in a DSSS system naturally implies a higher chip rate, which provides finer resolution of the multipath. This finer resolution means that the statistical character of the multipath on each resolved component will differ from that of a narrowband system, although it is not completely clear without further measurements as to what the exact manifestation of this difference will be [1]. Indications are that the Rayleigh model now used in so many current analyses will no longer be valid. Finer resolution will mean that each resolvable component will be composed of a smaller number of rays, which will result in less severe amplitude fading, but with a phase component that is somewhat random. With coherent RAKE reception (e.g., BPSK spreading with BPSK data modulation), each significant resolved component must

be tracked by a RAKE finger in order to contribute to the detection statistic. Of course, this tracking will be less than perfect due to the fading and noise in each finger. At some point, the additional multipath resolution may actually result in degraded receiver performance. It is the purpose of this paper to consider only the effect of imperfect phase tracking in the RAKE fingers on overall receiver performance. It will be shown that, for Rician amplitude fading and a Tikonov phase distribution in each finger, an optimum number of resolved paths actually exists. The Rician amplitude model for the RAKE finger fading components is chosen because it allows a tradeoff between specular and diffuse multipath components via the so-called K factor.

Analysis

Assuming BPSK modulation, and a RAKE receiver with maximal ratio combining and amplitude and phase error α_k and ϕ_k , respectively, on the k th finger, the detection statistic is [2]

$$U = 2E_b \sum_{k=1}^L |\alpha_k|^2 \cos \phi_k + \sum_{k=1}^L \alpha_k \cos \phi_k N_k \quad (1)$$

where N_k is the noise in the k th finger (assumed to be the same on all fingers) and E_b is the signal energy. Under the assumptions of this paper, α_k is Rician and ϕ_k is Tikonov with probability density function (pdf)

$$p(\phi_k) = \frac{\exp(\bar{\gamma}_{k, \text{loop}} \cos \phi_k)}{2\pi I_0(\bar{\gamma}_{k, \text{loop}})}, |\phi_k| \leq \pi \quad (2)$$

where $I_0(\cdot)$ is the modified Bessel function of order 0, $\bar{\gamma}_{k, \text{loop}} = \bar{\gamma}_k R_b / B_L$, R_b is the data rate, B_L is the loop bandwidth, and $\bar{\gamma}_k$ is the average signal-to-noise ratio (SNR) in the k th finger [i.e., $E(\alpha_k^2)$]. If the α_k 's are Rician distributed, then $\gamma_k = |\alpha_k|^2$ has the probability density function (pdf)

$$f_{\gamma_k}(z_k) = \frac{\exp\left[-\left(\frac{z_k}{\bar{\gamma}_k} + K\right)\right] I_0\left(2\sqrt{\frac{Kz_k}{\bar{\gamma}_k}}\right)}{\bar{\gamma}_k}, \quad z_k \geq 0 \quad (3)$$

The analysis uses the Gauss-Chebyshev quadrature method of [3] to obtain the probability of error from the moment generating function of the test statistic (1). Since (1) is conditionally Gaussian, given α_k and ϕ_k , the conditional moment generating function is easily written down. It may then be averaged with respect to the $\gamma_k = |\alpha_k|^2$ variables by straight forward integration. Averaging over the Tikonov-distributed ϕ_k 's takes a numerical integration for each value of the moment generating function variable $s = c(1 + j\tau_k)$ during application of the Gauss-Chebyshev quadrature rule for evaluation of the probability of error. Thus, using the fact that U is conditionally Gaussian, we write its moment generating function as

$$\Phi(s) = E[\exp(-sU)] = E_{\gamma_k, \phi_k} \left[\exp\left(-sm_U + s^2 \sigma_U^2 / 2\right) \right] \quad (4)$$

where the subscripts γ_k, ϕ_k mean that the expectation must still be done with respect to these random variables. Following [2], the required moments for the Gaussian moment generating function inside the right-hand expectation are

$$m_U = 2E_b \sum_{k=1}^L \gamma_k \cos \phi_k \quad \text{and} \quad \sigma_U^2 = 2E_b N_0 \sum_{k=1}^L \gamma_k \cos^2 \phi_k \quad (5)$$

where E_b is the bit energy and N_0 is the noise power spectral density. Using these moments, the moment

generating function becomes

$$\Phi(s) = \prod_{k=1}^L E_{\gamma_k, \phi_k} [\Phi_k(s | \gamma_k, \phi_k)] \quad (6)$$

where

$$\Phi_k(s | \gamma_k, \phi_k) = \exp[-2E_b \gamma_k \cos \phi_k (1 - 0.5N_0 \cos \phi_k s)s] \quad (7)$$

The expectation of (7) with respect to γ_k may be carried out by making the integrand of the expectation look like the integral of a Rician pdf, giving 1 since it is a pdf, with additional factors outside the integral. The result is

$$\Phi_k(s | \phi_k) = \frac{\exp[-KB\bar{\gamma}_k/(1+B\bar{\gamma}_k)]}{1+B}, \quad (8)$$

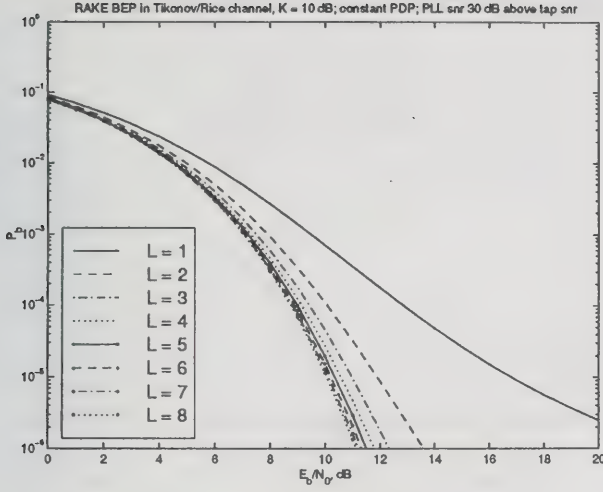
$$B = 2E_b \cos \phi_k s (1 - 0.5N_0 \cos \phi_k s)$$

The expectation with respect to ϕ_k was carried out numerically, and the probability of error was done using the Gauss-Chebyshev quadrature integration technique discussed in [3].

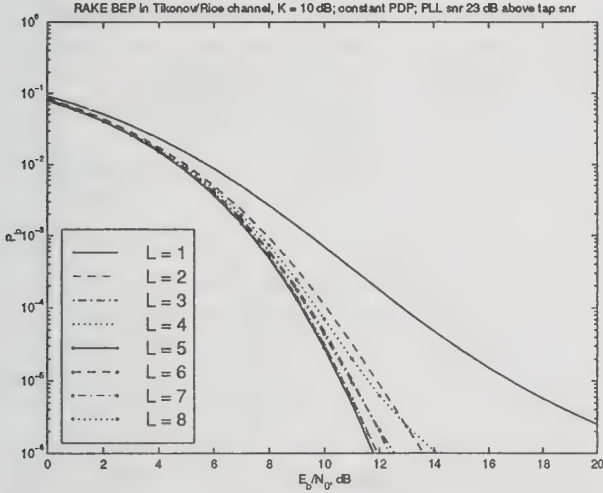
Results

The procedure describe above was used to analyze BPSK RAKE system performance in delay spread multipath with Tikonov phase error in each finger. Typical results showing bit error probability, P_b , versus E_b/N_0 are given in Fig. 1 for a constant power delay profile and the Rician specular-to-diffuse ratio $K = 10$ dB parameterized on the order of diversity, L . Figure 2 shows corresponding results for $K = 0$ dB. Note that for a sufficiently low bit rate to loop bandwidth ratio and sufficiently high E_b/N_0 there is first improvement with increasing L but eventually performance degrades. This indicates that as signal-to-noise ratio decreases for each diversity path (finger in the RAKE receiver), the phase tracking jitter manifested through the Tikonov pdf overcomes any improvement due to additional diversity. This behavior is better illustrated by Figures 3 and 4 which show P_b versus L for a fixed E_b/N_0 of 10 dB for the cases of $K = 10$ and 0 dB, respectively.

The behavior that was expected is observed in that for sufficiently high loop SNR of the RAKE fingers,



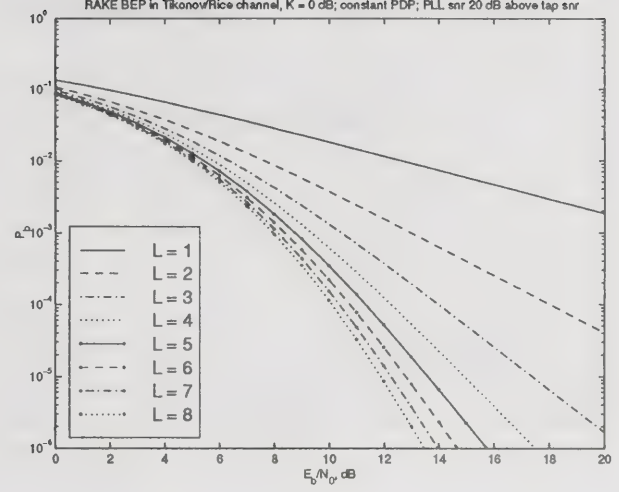
(a)



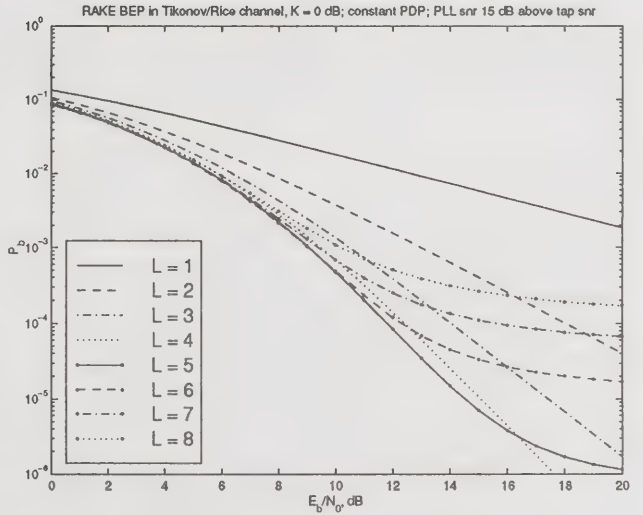
(b)

Figure 1. Bit error probability versus E_b/N_0 in Rician fading with $K = 10$ dB with L as a parameter: (a) loop snr 30 dB above finger snr; (b) 23 dB above finger snr.

use of additional diversity, L , gives added improvement whereas, for sufficiently low loop SNR, improvement is first observed with increasing L with an eventual bottoming out, and then performance gets worse. In the case of constant power delay profile as illustrated here, this behavior is dramatic since all tracked multipath components have equal power. In the case of an exponential power delay profile, the change versus L is much less dramatic because additional tracked multipath components have exponentially less power than those already tracked.

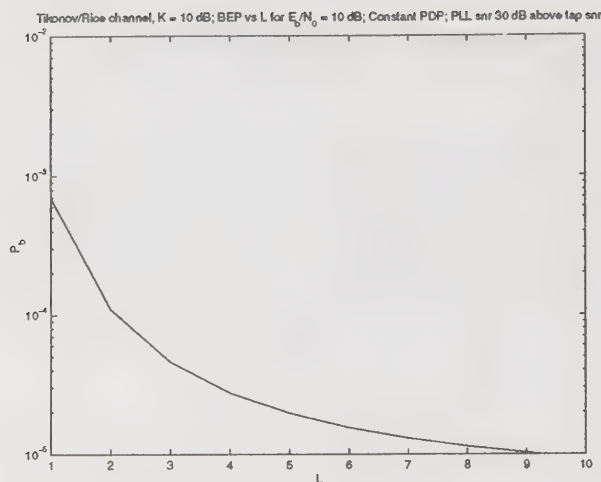


(a)

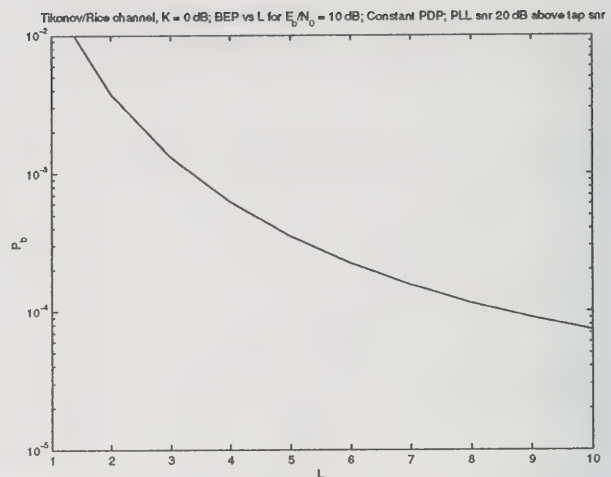


(b)

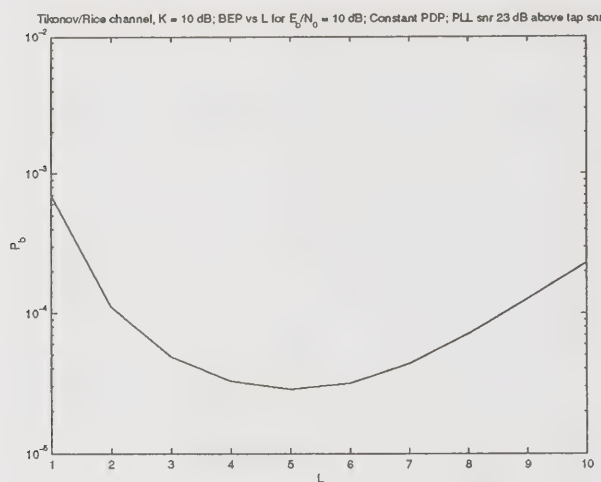
Figure 2. Bit error probability versus E_b/N_0 in Rician fading with $K = 0$ dB: (a) loop snr 20 dB above finger snr; (b) 15 dB above finger snr ($K = -\infty$ dB is Rayleigh).



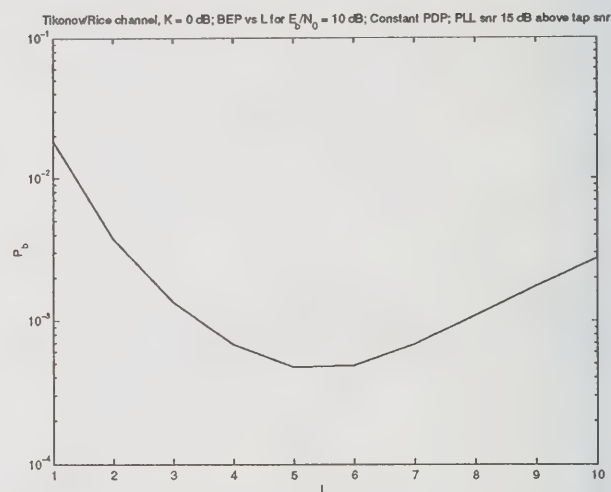
(a)



(a)



(b)



(b)

Figure 3. Bit error probability versus L in Rician fading with $K = 10$ dB with $E_b/N_0 = 10$ dB: (a) loop snr 30 dB above finger snr; (b) 23 dB above finger snr.

Figure 4. Bit error probability versus L in Rician fading with $K = 0$ dB with $E_b/N_0 = 10$ dB: (a) loop snr 20 dB above finger snr; (b) 15 dB above finger snr.

Conclusion

A numerical method has been summarized for analyzing the effect of phase estimation in the fingers of a coherent RAKE receiver. The model used assumes Rician fading amplitudes and Tikonov phase pdf in the fingers. Only results for constant power delay profiles were presented due to space limitations. Similar results for exponential power delay profiles show that the minimum versus multipath components tracked is much less dramatic due to the decrease in power with each component tracked. The effect of estimation of other parameters, such as code timing and symbol synchronization, is currently being examined.

References

- [1] M. Win & R. Scholtz, "On the robustness of ultra-wideband signals in dense multipath environments," *IEEE Commun. Letters*, Vol. 2, pp. 51-53, Feb. 1998.
- [2] J. G. Proakis, *Digital Communications*, 3rd ed., New York: McGraw-Hill, 1995.
- [3] E. Biglieri, et al., "Simple method for evaluating error probabilities," *Electr. Letters*, Vol. 32, pp. 191-192, Feb. 1, 1996.

Use of Prony's Method For Extracting The Poles And Zeros Yielding A Wideband Window Type Response Of Circular Antenna Arrays

Assoc. Prof. Dr. Fatma M. El-Hefnawi, Senior Member, IEEE
Electronics Research Institute, Microwave Engineering Department
E-mail: Fatma@eri.sci.eg
Tel. (202) 2431297

Abstract

Most antennas or antenna arrays response are not needed at a single frequency, but rather a continuous representation over a specified period is necessary. Hence, the idea of extracting the poles in the complex plane so as to cover a wider response in frequency or time domain for the antenna circular array elements are calculated using rational function approximation in the complex frequency domain of double infinite series with unknown coefficients. The solution starts using electric field integral equation, application of the method of moments and extended boundary conditions. Then the poles are calculated using Prony's method for the input impedance and admittance of each antenna element. A newly calculated window type input impedance suffering from no singularities can be obtained. Impedances' or admittances' in the complex frequency plane are calculated by repeated application of the method of moments without the need of a newly computed matrix inversion. The method used in this paper is analogous to the formal method by the same author, which was done in the real frequency and without extracting the poles but rather using the inverse of the function describing the current, impedance or admittance. Also, the current distribution assumed was not a rational function in the complex frequency domain.

Key words: Prony's method, circular array.

Introduction

In computational electromagnetics, the method of moments[1] has been well established in antenna work especially with the extensive use of digital computer. Also, in the analysis and design of antennas and arrays, it is often required to generate a broad-band response which is often done by solving the problem of interest at a set of discrete frequencies. If the electromagnetic response is very sharp then it becomes quite inefficient regarding time and effort consumption, to repeat the computation process using the method of moments at each frequency in the required band. Hence, the need for analytic continuation to interpolate or extrapolate the computed narrow band results to cover a wider range. This is done in this paper by assuming the current over each element in the circular array (as an example) by a rational function of two polynomials in the complex frequency domain. This approximation is known as Pade approximation[2]. The objective will then be the determination of the unknown coefficients of the numerator and denominator polynomials of each element. The currents on each element is calculated using the electric field integral equation EFIE at a single frequency $s = s$. Singularities are extracted out by calculating the zero in the denominator polynomials of each element using Prony's method. The zero of the numerator polynomials is also calculated when the input

impedance is needed. Singularities are extracted out, and a new function is derived and calculated in the complex frequency domain. Then, the original function is retained back. The whole process needs no matrix calculation or inversion which numerically arose in the method of moments.

Integral Equation Formulation Of Circular Array In the Complex Frequency Domain:

The EFIE governing the circular array of P thin dipoles $a \ll \lambda$, length $2h$ in the complex frequency domain is as follows:

$$-s\epsilon_0 E_{inc_p}(z, s) = \sum_{m=1}^P \int_{-h}^{+h} I_m(z', s) K_{mp}(z, z', s) dz' \quad p=1,2,3,\dots,P \quad (1)$$

where $E_{inc_p}(z, s)$ is the incident electric field on the p^{th} element, $I_m(z', s)$ is the current distribution on the m^{th} element, $K_{mp}(z, z', s)$ is the kernel and is defined by :

$$K_{mp}(z, z', s) = \left(\frac{\partial^2}{\partial z^2} - \frac{S^2}{C^2} \right) \frac{e^{-\frac{s}{c} R_{mp}}}{R_{mp}} \quad (2)$$

where $R_{mp} = \sqrt{a^2 + (z - z')^2}$ if $m \neq p$

$$R_{mp} = \sqrt{b_{mp}^2 + (z - z')^2} \text{ if } m \neq p \quad (3)$$

$$\text{and } b_{mp} = 2\rho \sin\left[\frac{(p-m)\pi}{p}\right] \quad (4)$$

where ρ is the radius of the circular array, assuming that extended boundary condition is applied[3,4] to avoid singularities in the kernel, and that the elements in the circular array are evenly distributed. Using the same procedure used in [6] except $s = j\omega$. And applying the method of moments, choosing the same kernel expansions for mutual and self kernel and the same current expansion used by the author in [6]. Then Pade approximation for the current is applied. The incident exciting field $E_{inc}(z, s)$ is assumed as a frill of magnetic current which resembles the exciting field in the aperture of a ground plane where TEM modes are assumed, thus

$$E_{inc}(z, s) = \frac{V_p}{2 \ln(b/a)} \left(\frac{e^{-\frac{s}{c}R_a}}{R_a} - \frac{e^{-\frac{s}{c}R_b}}{R_b} \right) \quad (5)$$

$$\text{where } R_a = \sqrt{a^2 + z^2}, R_b = \sqrt{b^2 + z^2} \quad (6)$$

Computation Of Current Expansion And It's Derivatives In The Complex Frequency Domain :

The current distribution on the p^{th} element given in equation (1) is expanded making use of lade Cauchy's method[5], such that the boundary condition at the end of the wave is satisfied i.e. $I(\pm h) = 0$. Thus,

$$I_p(z', s) = \frac{A_p(s)}{B_p(s)} \quad (7)$$

$$\text{where } A_p(s) = \sum_{\alpha=0}^x a_{p\alpha} s^\alpha \quad (8)$$

$$B_p(s) = \sum_{\beta=0}^y b_{p\beta} s^\beta \quad (9)$$

$$\text{and } b_{p0} = 1 \quad (10)$$

with out loss of generality. Equation (7) can be written as:

$$A_p(s) = I_p(z', s) B_p(s) \quad (11)$$

Differentiating equation (11) n times with respect to s gives:

$$A_p^{(n)}(s) = \sum_{i=1}^n {}^n C_i I_p^{(n-i)}(z', s) B_p^i(s) \quad (12)$$

$$\text{and } A_p^{(n)}(s) = \sum_{i=1}^p \frac{i!}{(i-n)!} a_i s^{i-n} \quad (13)$$

The number of unknown in the current expression is equal to $x+y+1$. Equation (7) needs $(x+y+1)P$ points. This can be overcome by calculating the current at one or two frequencies then the frequency derivatives are calculated also such that the total number of the current values and frequency should be equal to $(x+y+1)P$. The current derivatives are calculated from the following equation which normally arose as follows:

$$[V] = [Z][I] \quad (14)$$

$$\text{or } [I] = [Z]^{-1}[V] \quad (15)$$

where, $[I]$ is a vector of sub-vectors representing the current on each element of the circular array, $[V]$ is the known excitation to each antenna element and is also composed of sub-vectors for each element, and $[Z]$ is the matrix that describes the interaction between the different elements in the array. Equation (15), when differentiated n times with respect to complex frequency s gives:

$$I^{(n)}(s) = Z^{-1}(s) \left[V^{(n)}(s) - \sum_{i=1}^{n-1} {}^n C_i [Z_n]^{(n-i)} I^{(i)}(s) \right] \quad (16)$$

where the (n) represents the n^{th} derivative with respect to frequency s , ${}^n C_i$ is the binomial factorial given by:

$${}^n C_i = \frac{n!}{(n-i)! i!} \quad (17)$$

Equation (15) and (16) are solved to get $a'_p s$, $b'_p s$ appearing in equation (8), (9). Thus one gets the current distribution as a function of (s) for each antenna in the array.

Extraction of Poles Using Prony's Method For The Current Distribution:

It is well known that the current on each element of the circular array suffers from singularities. Also the zero in the current are poles of the input impedance. Thus one has to calculate both zeros and poles using Prony's method[7].

A new source currents are then calculated suffering from no singularities using the following equation:

$$i(z', s) = (s - s_{pole}) I(z', s) \quad (18)$$

Equation (18) when differentiated n times with respect to frequency s gives:

$$i''(z', s) = (s - s_{pole}) I''(z', s) + n I^{n-1}(z', s) \quad (19)$$

Equation (19) is used to get $i''(z', s)$ suffering from no singularities using the formally calculated current derivatives $I''(z', s)$. Hence a new function is derived and calculated for different values of (s) . Thus,

$$i_p(z', s) = \frac{\sum a_{new} s^\alpha}{\sum b_{new} s^\beta}, \quad p=1,2,3,\dots,P \quad (20)$$

where new values for a 's and b 's are calculated from (14). Then $I(z', s)$ are calculated back again using the following equation:

$$I_p(z', s) = i_p(z', s) / (s - s_{pole}), \quad p=1,2,\dots,P \quad (21)$$

Results and Conclusions

The circular array chosen is a four element circular array of thin dipoles with one element driven. The current expansion used in equation (7) is in the form of rational functions where the numerator polynomial has a higher order than the denominator for each antenna element in the circular array. Only one calculation for the value of the current is computed using the method of moments. Frequency derivatives are used instead of many sample frequency values for the current. The number of frequency derivatives for the current on each element needed for the solution is changed from four derivatives to eight where the results are fairly good and no need for more derivatives. Although up to thirteen derivatives are calculated and the results are checked for stability. Results when extracting the first pole in the denominator are plotted, which starts to diverge when it starts to become near the second pole ($s = 0.54$). A newly band of results can be obtained if a new frequency is chosen after the appearance of the first pole and a newly calculated frequency derivatives can be calculated and the process for extracting a new pole can be done so as to cover the new band of frequencies till the second pole. It is to be noticed that the poles can be readily calculated without the need of calculating the frequency derivatives from the zeros of the denominator polynomial in equation (7).

Results for self and mutual conductances and susceptances G_{11} , B_{11} , G_{12} , B_{12} , G_{13} , B_{13} are plotted in figures [1-2] and are compared with experimental results of R. B. Mack[8] and the results for the author in [9] with satisfactory agreement.

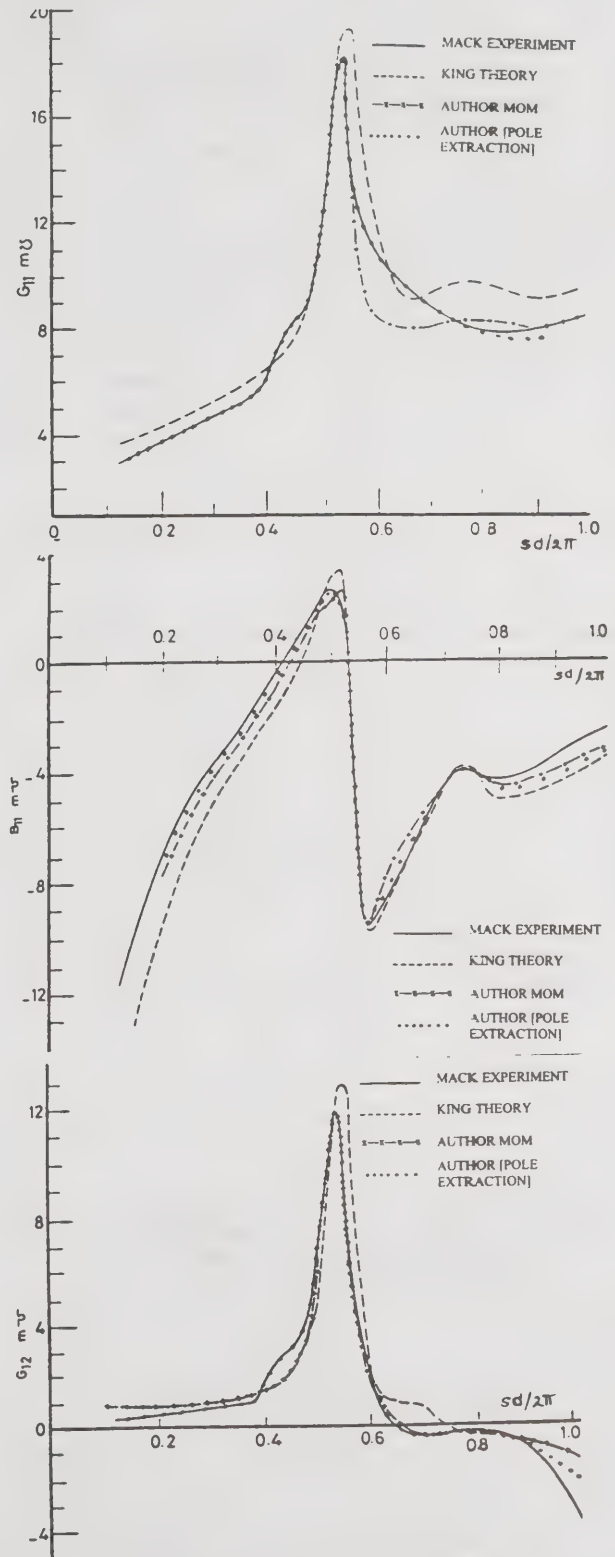


Fig.(1) Mutual Conductances and Susceptance of four Element Circular Array G_{11} , B_{11} , and G_{12} (d =dipole spacing, $sh=j0.50PI$, $sa=j0.01444PI$)

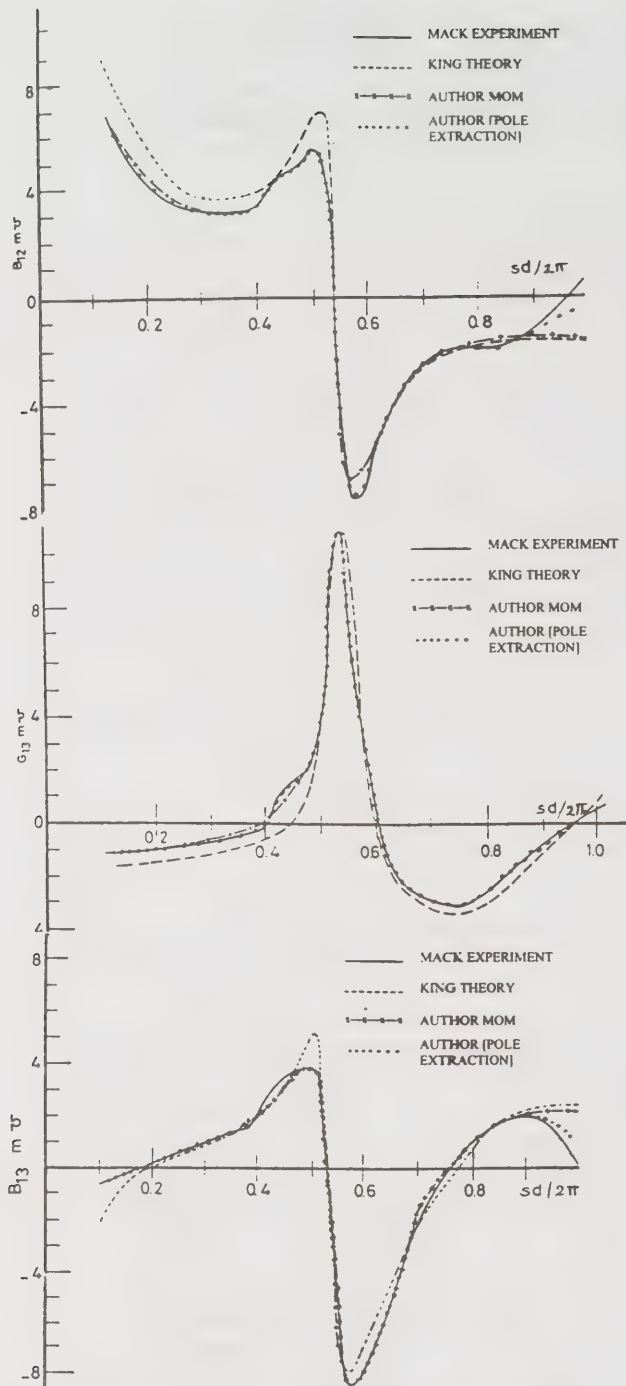


Fig.(2) Mutual Conductance and Susceptances of four Element Circular Array B12,G13, and B13
(d =dipole spacing, $sh=-j0.50P1$, $sa=-j0.01444P1$)

References

- [1] R.F. Harrington, Field Computation by moment methods, New York, Macmillan, 1968.
- [2] R.W. Hamming, "Numerical methods for scientists and engineers", New York, Me. Gram-Hill, 1962.
- [3] K.A. Al-Baidwaihy and J.L. Yen, "Extended boundary condition integral equation for perfectly conducting and dielectric bodies: formulation and propagation, VVP. AP-22, No.4, PP 546-551, July 1975.
- [4] W.L. Stutzman, G.A. Thiele, "Antenna theory and design", New York, John-Wiley, 1998.
- [5] K. Kottapalli *et al*, "Accurate computation of wide-band response of electromagnetic system utilizing narrow-band information", IEEE Trans Antennas and Propagation, Vol. 39, No.4, PP. 632-687, April 1991.
- [6] F.M. El-Hefnawi, "Frequency perturbation for a circular array of coupled cylindrical dipole antennas", Microwave and optical technology letters, Vol. 12, No. 1, PP. 36-40, May 1996.
- [7] G.R.B. Prony, "Essai experimental et analytique sur les lois de la dilatalrlite de fluides elastiques et sur cells de la vapeur de l'alcool, à différents tempoeatures", Journal de l'Ecole Polythnique (Paris), Vol. 1, No. 2, PP. 24-76, 1795.
- [8] F. M. EL-Hefnawi, "Analysis and Numerical Solution of the Cylindrical Dipole Antenna and Circular Array", Ph.D. Thesis, Cairo University, 1979.
- [9] R. B. Mack, "A study of Circular Arrays", Ph.D. Thesis Harvard University, 1969.

Microstrip Bowtie Patch Antenna for Wireless Indoor Communications

C. H. Ng, S. Uysal and M. S. Leong
Department of Electrical Engineering,
National University of Singapore,
10 Kent Ridge Crescent,
Singapore 119260

+65-775 6666, +65-779 1103 (fax), clesu@nus.edu.sg

Abstract

A novel methodology is presented for the design of a wideband microstrip bowtie patch antenna (MBA) of finite length placed on a dielectric substrate. Here, a microstrip line feed network was used in HP Microwave Design System (MDS™) to match the MBA, since this type of feed network facilitates the construction of an array. Experimental measurements for the bowtie designed in the X-band are shown to agree well with the simulated predictions, especially the radiation patterns. The resonant length of the antenna was found to be around twice the dielectric wavelength of 20.2mm. The bandwidth obtained is about 9.7%.

Introduction

The major limitation of microstrip antennas is the narrow bandwidth. Defining bandwidth as the frequency range over which SWR is 2:1 or less, patch antenna bandwidths are typically about 1% to 5%. Thus, considerable effort has been focused on developing antennas for wideband operation. One such antenna is the bowtie antenna, known for its simple design and broadband performance.

This paper presents the microstrip bowtie patch antenna for wireless indoor communications, fabricated on a Taconic ($\epsilon_r = 2.2$, $h = 0.79$ mm) substrate. Its S11 response and radiation patterns are also given in this paper.

Design of Microstrip Bowtie Antenna

Following the design procedure of [1], where matching is accomplished with HP Microwave Design System (MDS™), the circuit is optimised towards a goal of $S_{11} < -15$ dB for the frequency range of 9.65GHz to 10.35GHz. The S11 response shows a very large dip of about -45dB at a resonant frequency of 9.74GHz and the bandwidth obtained is a significant 8.8%. The various parameters in the MBA design are illustrated in Figure 1 and its values are given in Table 1.

The MBA is simulated in ENSEMBLE™ from Boulder Microwave Technologies Inc. The simulated S11 response shown in Fig. 2 gives a dip of -40.26dB at the resonant frequency of 9.94GHz with a bandwidth of

9.26%. This set of parameter values are adopted for fabrication and testing.

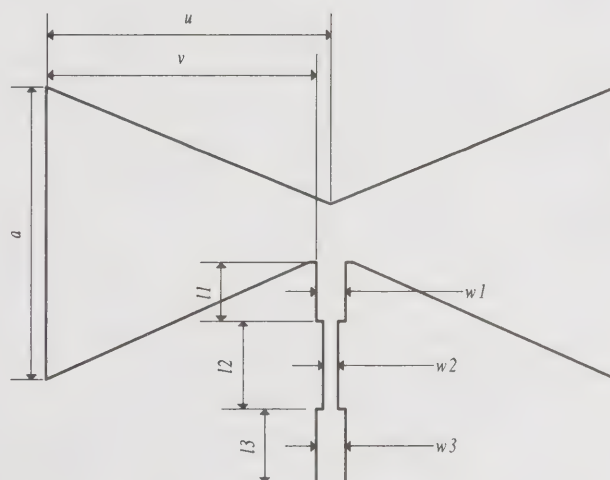


Figure 1: Layout of MBA matched to 50Ω coaxial feed

Table 1: Design values for MBA

Parameter	Dimension (mm)
$l1$	4.26
$l2$	6.87
$l3$	5.54
$w1$	1.88
$w2$	0.91
$w3$	2.40
a	23.97
c	5.00
u	20.76
v	19.56

Results and Discussions

A. Simulated and Measured S11 Response

The measured return loss for MBA is compared with the ENSEMBLE™ simulated results in Figure 2. The simulation was done from 9.5 - 10.5GHz to give emphasis to the bandwidth over the resonant frequency, whereas the measurement was made over the X-band. The simulated response shows a bandwidth of 9.26% whereas the measured bandwidth gives 9.71%, showing the desired wideband characteristics. The shape of the simulated S11 response is similar to the measured response, with the later recording a smaller dip of -33.31dB at the resonant frequency of 10.2GHz, compared to a dip of -40.26dB at 9.94GHz for the simulated response. Two dips were also recorded at 9.86GHz (-23.87dB) and 10.56GHz (-16.66dB).

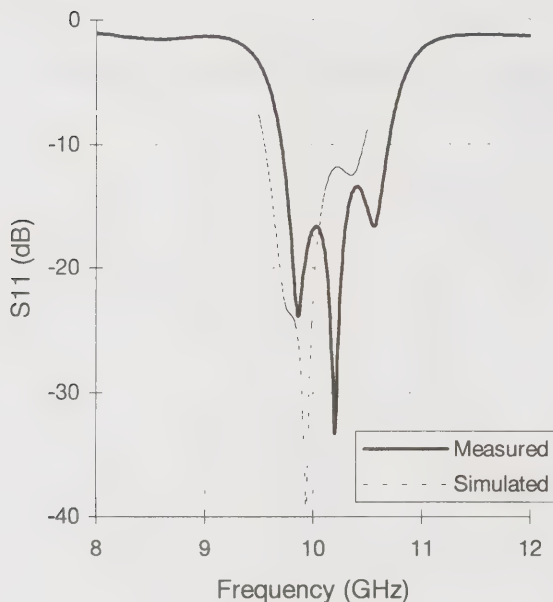


Figure 2: Simulated and measured return loss for MBA on Taconic substrate of $\epsilon_r = 2.2$ and $h = 0.79\text{mm}$

B. Simulated and Measured Radiation Pattern

Similarly, the radiation patterns for MBA was simulated in ENSEMBLE™ over 9.5 - 10.5GHz. The simulated patterns show close resemblance to the measured patterns. The simulated and measured patterns are as shown in Figures 3 - 5 for the resonant frequency of 10.2GHz.

It can be seen that the E_θ and E_ϕ components for both simulated and measured patterns are symmetrical about

the 0° axis. At 10.2GHz, the E_θ component shows a major dip in the broadside direction ($\theta = 0^\circ$) and a relatively minor dip at $\approx \pm 45^\circ$. The maximum power ratio for this component is at $\pm 19.35^\circ$. The E_ϕ component recorded its maximum power ratio at $\pm 29.25^\circ$ with a very slight dip at broadside.

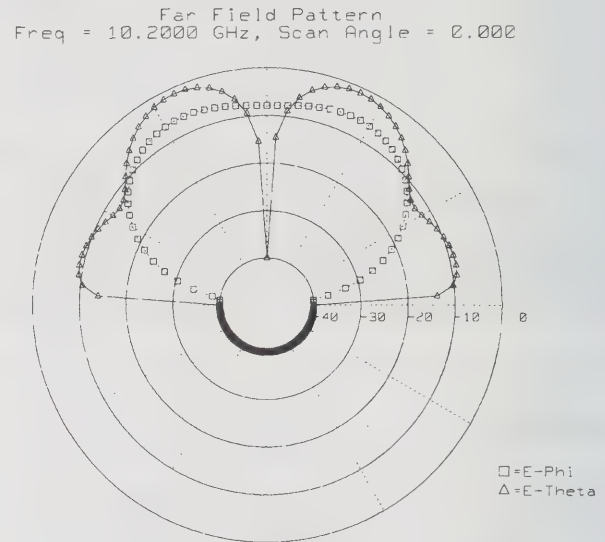


Figure 3: Simulated radiation pattern at resonant frequency of 10.2GHz

Conclusions

A modified MBA is presented and the broadband matching technique using microstrip lines in MDS™ was used to match the MBA to a 50Ω coaxial feed. The shapes of the measured S11 response and radiation patterns show good agreement with the simulated results, particularly the latter. The measured S11 response gives the wideband characteristic of the antenna to be 9.7%. The measured radiation patterns also show well defined lobes in both the E and H planes.

Reference

- [1] Loi K. W., S Uysal and M. S. Leong, "Novel Design of a wideband microstrip bowtie patch antenna", IEE Proceedings, H, Microwaves, Antennas and Propagation. April 1998 (United Kingdom).

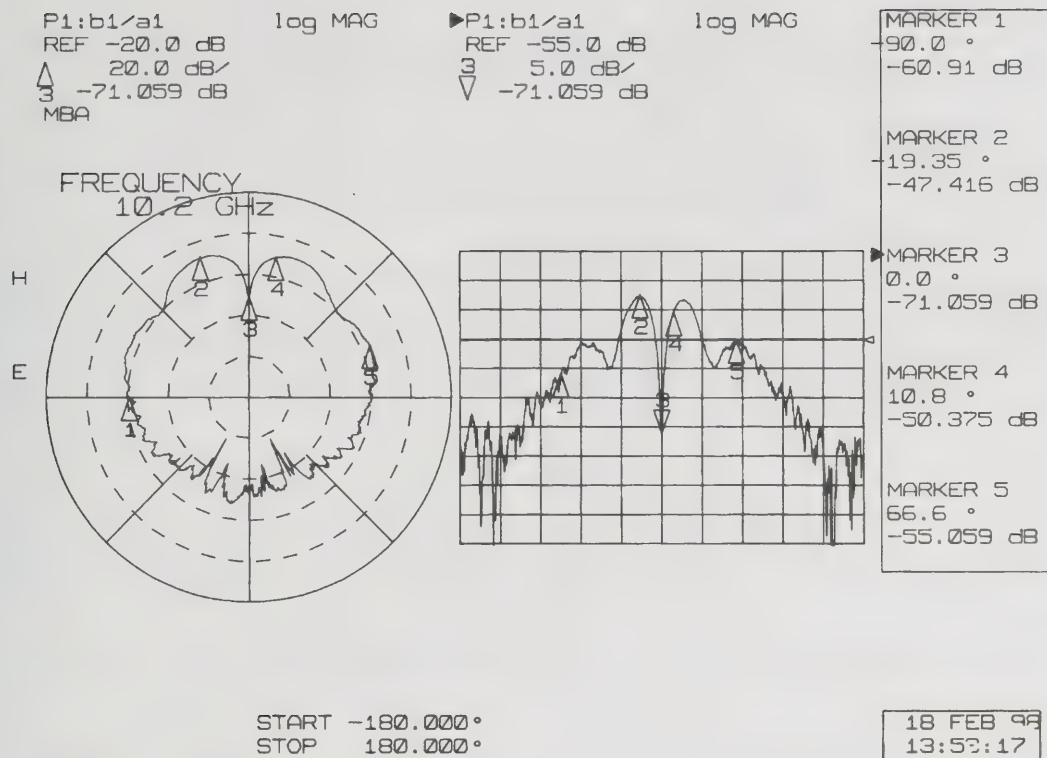


Figure 4: Measured radiation pattern for E_θ component at resonant frequency of 10.2GHz

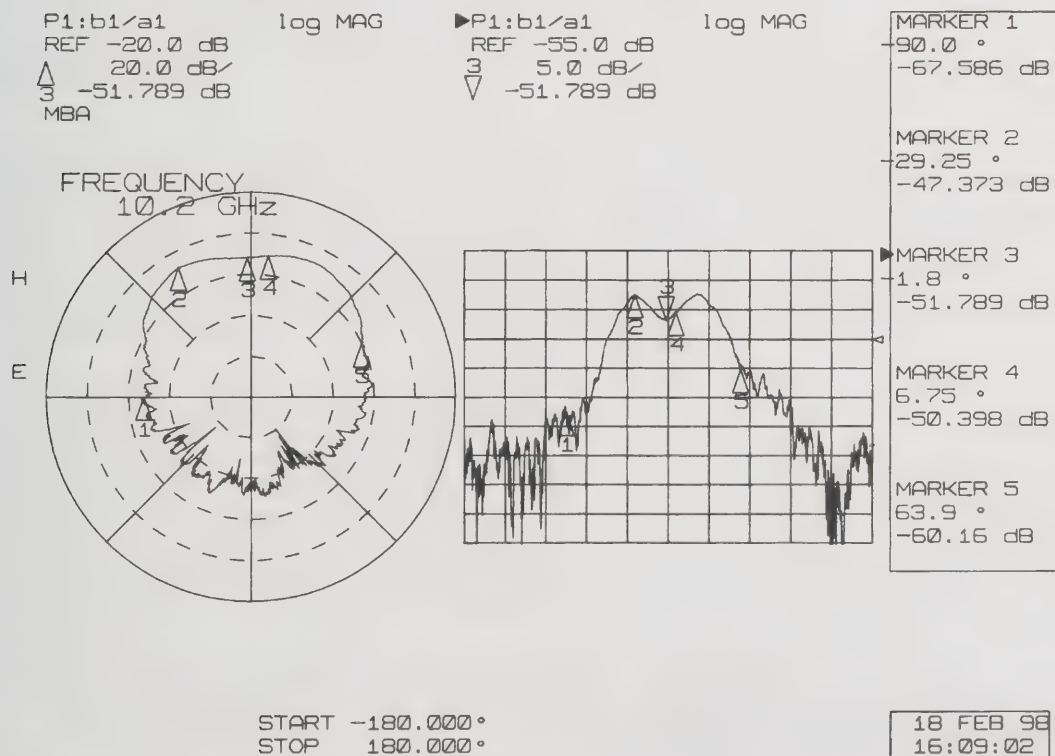


Figure 5: Measured radiation pattern for E_θ component at resonant frequency of 10.2GHz

Beamforming applied to an Adaptive Planar Array

Ana Maria G. Guerreiro, Adrião Duarte D. Neto, Fábio Adriano Lisboa
Rio Grande do Norte Federal University
Technology Center
Eletrical Engineering Department
Computation and Automation Engineering Laboratory
59.072-970 Campus Universitário - Lagoa Nova - Natal - RN -Brazil
anamaria@leca.ufrn.br, adriao@leca.ufrn.br, fabio@leca.ufrn.br

Abstract

This paper presents algorithm to an adaptive array beamforming. The adaptive process will be developed through a neural linear network based on the complex Least Mean Square (LMS) algorithm. This study is developed for a planar antenna arrays.

Key words: planar array, beamforming, generalized sidelobe canceler, adaptive algorithm

Introduction

Adaptive arrays beamforming has been widely used in many areas, in a variety of applications, radar, sonar, communication systems, geophysical exploration and other systems that are required to reject interferences. For all the systems above mentioned, the general goal is to reconstruct of a desired signal in the presence of interferences and noise.

Beamforming is a technique to receive a desired signal from a known direction of arrival and reject interferences by controlling the beam pattern of an antenna array in an adaptive way. Many researches have been developed in this area [1]-[5], like the works developed by Frost [6], Griffiths [7]-[8] and more recent researches [9]-[10].

A special class of beamformers are those that filters the signal from a specified direction of interest or desired and considers all other signals as interference or noise. In this class, we have particular interest in the Generalized Sidelobe Canceler (GSC) structure. The GSC employs a linearly constraint on the response of the beamformer so that signals from the direction the interest are passed with specified gain and satisfied two basic conditions:

- (a) The imposition of linear multiple constraints designed to preserve an incident signal along a direction of interest.
- (b) The adjustment of some weights, in accordance with the LMS algorithm, minimize the effects of interference and noise at the beamformer output.

The previous works consider linear antenna array [11]-[12]. In this paper we consider for our research the planar antenna that has disadvantages as low radiation efficiency, small bandwidth and low gain so adaptive process is going to be a suitable technique for those antennas.

Generalized Sidelobe Canceler

For array beamforming, we consider of a planar antenna array of $N_x \times N_y$ elements. They are distributed uniformly spaced at known spatial locations. The beamformer output is defined by the inner product:

$$y(n) = \mathbf{w}^H(n) \mathbf{u}(n) \quad (1)$$

$$\mathbf{w}(n) = [w_1, w_2, w_3, \dots, w_M] \quad (2)$$

where $\mathbf{w}(n)$ is a complex weights vector, $M = N_x \times N_y$ is the numbers of antennas elements of the planar array and $\mathbf{u}_i(n)$ is the individual response of element i and it has the form [2]:

$$\mathbf{u}_i(n) = [u_0(n), u_1(n)e^{-j\xi}, \dots, u_{M-1}(n)e^{-j(M-1)\xi}]^T \quad (3)$$

The important thing to notice that in our work the incoming signals depend on the spatial direction (θ, ϕ) . It is define in x-y-z plane as shown in figure 1. The $\beta_i(\theta, \phi)$ is the relative time delay of this element to an arbitrarily chosen spatial reference point. The time delay of element i is given by:

$$\beta_i(\theta, \phi) = \frac{(x_i \cos \phi \sin \theta + y_i \sin \phi \sin \theta + z_i \cos \theta)}{c} \quad (4)$$

$$\xi = 2\pi f \beta_i(\theta, \phi) \quad (5)$$

where c is the propagation speed of the incoming signals.

We assume that the element spacing of the array was chosen as one half of the wavelength, so specifically, $\beta_i(\theta, \phi)$ has the form:

$$\beta_i(\theta, \phi) = \pi \cos \phi \sin \theta + \pi \sin \phi \sin \theta \quad (6)$$

The optimum weight vector is found by minimizing the array output power subject to K linear constraints as follows:

Minimize:

$$E[|y(n)|^2] = \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (7)$$

$$\mathbf{R} = E[\mathbf{u}(n)\mathbf{u}(n)^H] \quad (8)$$

Subject:

$$\mathbf{C}^H \mathbf{w} = \mathbf{g} \quad (9)$$

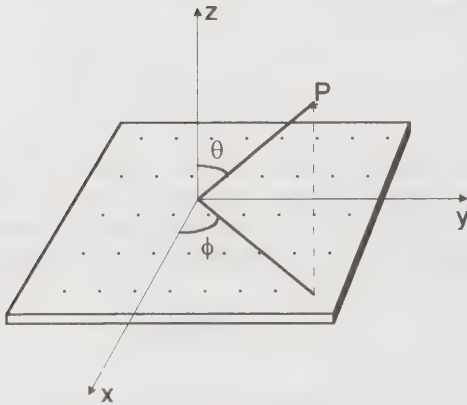


Fig. 1: Planar antennas array.

where H denotes conjugate transpose and the correlation matrix \mathbf{R} , does not require measurements of the correlation functions, the LMS algorithm represents an estimate whose expected value approaches the Wiener solution as the number of iterations approaches infinity.

The matrix \mathbf{C} is the constraint matrix, and \mathbf{g} is the gain vector, has constant elements. Assuming K linear constraints, the matrix \mathbf{C} is an $(N_x * N_y) \times K$, and \mathbf{g} is an $K \times 1$ vector; the rows number of \mathbf{C} is defined with the number of antennas elements. Furthermore, we assume columns.

For example:

$$[\mathbf{s}(\phi_0, \theta_0), \mathbf{s}(\phi_1, \theta_1)]^H \mathbf{w} = [1 \ 0]^T \quad (10)$$

$$\mathbf{s}(\phi, \theta) = [1, s_1(n)e^{-j\xi}, \dots, s_{M-1}(n)e^{-j(M-1)\xi}]^T \quad (11)$$

where $\mathbf{s}(\theta_0, \phi_0)$ is the steering vector that preserve the desired signal on the planar array along the direction (θ_0, ϕ_0) and suppress an interference know to originate along the (θ_1, ϕ_1) .

The orthogonal complement of a subspace \mathbf{C} contains every vector that is perpendicular to \mathbf{C} [13]. It is important to noticed that the four fundamentals linear algebra subspaces are more than just orthogonal, they are orthogonal complements. In another way, the nullspace of \mathbf{C} is the complement orthogonal of the row space and the nullspace of \mathbf{C}^T is the complement orthogonal of the column space. The columns of a matrix $\mathbf{C}\mathbf{a}$ is defined as a basis for the orthogonal complement of space spanned by the columns of matrix \mathbf{C} , so the matrix $\mathbf{C}\mathbf{a}$ is the nullspace of the matrix \mathbf{C}^T . The matrix $\mathbf{C}\mathbf{a}$ is called the signal-blocking matrix \mathbf{C} . Using the definition of an orthogonal complement, we may write:

$$\mathbf{C}^H \mathbf{C}\mathbf{a} = 0 \quad (12)$$

or

$$\mathbf{C}\mathbf{a}^H \mathbf{C} = 0 \quad (13)$$

Next, let define a nonadaptive beamformer component represented by:

$$\mathbf{w}_q = \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{g} \quad (14)$$

We may define express the overall weight vector of the GSC structure as shown in the figure 2 as:

$$\mathbf{w} = \mathbf{w}_q - \mathbf{C}\mathbf{a}\mathbf{w}_a \quad (15)$$

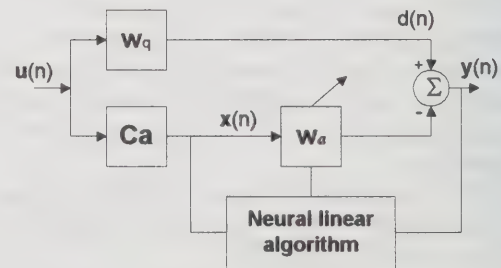


Fig. 2: Block diagram of GSC.

The equation (14) shows that the weight vector \mathbf{w}_q is that part of the weight vector \mathbf{w} that satisfies the constraints. In contrast \mathbf{w}_a is unaffected by the constraints. Substituting equation (15) in (1):

$$y(n) = \mathbf{w}_q^H \mathbf{u}(n) - \mathbf{w}_a^H(n) \mathbf{C} \mathbf{a}^H \mathbf{u}(n) \quad (16)$$

$$y(n) = d(n) - \mathbf{x}(n) \quad (17)$$

We may rewrite equation (16) as:

$$y(n) = d(n) - \mathbf{w}_a^H(n) \mathbf{x}(n) \quad (18)$$

$d(n) \rightarrow$ Desired response of GSC

$\mathbf{x}(n) \rightarrow$ Input vector

Adaptive Algorithm

The linear adaptive filtering algorithm consists of two basic process [11]:

1. A filtering process, which involves: computing the output of the filter and generating an estimation error by comparing this output to a desired response.
2. An adaptive process, which involves the automatic adjustment of the tap weights in accordance with the estimation error.

Summary of the algorithm:

Parameters:

- The desired and interference angles incident
- The number of antennas elements = $N_x \times N_y$
- Step-size parameter = η

- 1) Initialization of weights.
- 2) Compute the nonadaptive weights, \mathbf{w}_q with equation (14).
- 3) Compute the signal- blocking matrix, $\mathbf{C} \mathbf{a}$ that is the null space of the matrix \mathbf{C}^T .

Computation $n = 0, 1, 2, 3, \dots$, compute

$$y(n) = d(n) - \hat{\mathbf{w}}_a^H(n) \mathbf{x}(n) \quad (19)$$

$$\hat{\mathbf{w}}_a(n+1) = \hat{\mathbf{w}}_a(n) + \eta y(n) \mathbf{x}^*(n) \quad (20)$$

Results

For the experiment we consider the linear neural network based on the complex LMS algorithm applied to an adaptive Generalized Sidelobe Canceler beamformer consisting of a planar arrays of $N_x \times N_y$ uniformly spaced antennas elements. The element

spacing of the array was chosen as one half of the wavelength. There are two components: a target signal in the direction of interest (θ_0, ϕ_0) , and interference in another direction (θ_i, ϕ_i) . It's assumed that these two components originated from independent sources.

The aim of the simulations is to investigate the evolution of the adaptive beamformer response for varying interference-to-noise ratio, varying number of iterations and to evaluate the effect of varying the target-to-interference ratio on the interference-nulling performance of the beamformer.

The figure 3 and 4 show the conventional uniform array patterns for 4x4 and 3x3 elements planar arrays. The figure 5 (4x4 planar array) and 6 (3x3 planar array) observes the nulls steered to direction of arrival of the interferences and the capability of the beamformer improves with increasing number of iterations. As a result, the interferences are suppressed improving the performance of the adaptive beamformer.

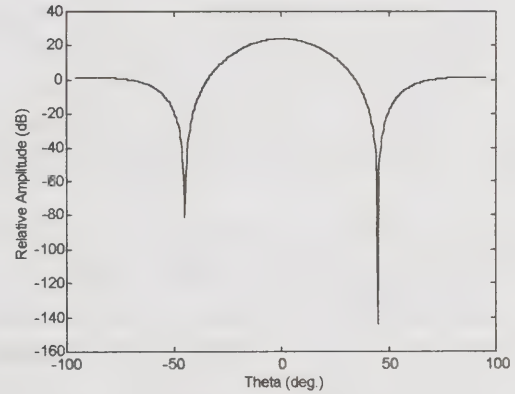


Fig. 3: The conventional uniform array pattern for 4x4 element planar array. The look direction angles are $\phi_0 = 30^\circ$, $\theta_0 = 60^\circ$.

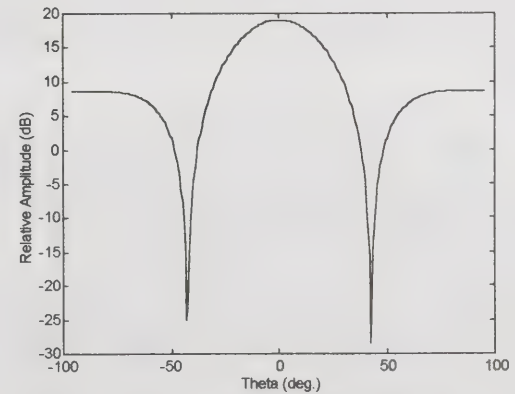


Fig. 4: The conventional uniform array pattern for 3x3 element planar array. The look direction angles are $\phi_0 = 20^\circ$, $\theta_0 = 30^\circ$.

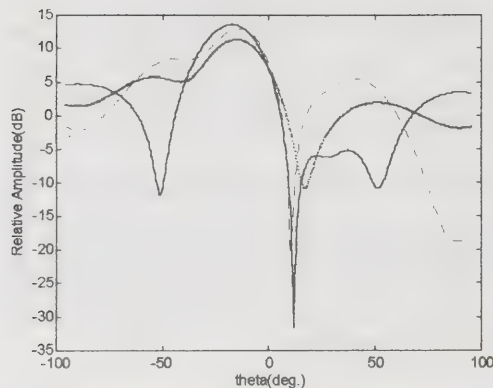


Fig. 5: Adapted response for SNR=20dB, with the null steered to direction ($\phi = 45^\circ$, $\theta = 10^\circ$), varying number of iterations. Solid curve=20, dashed curve=30 and point curve=10 iterations.

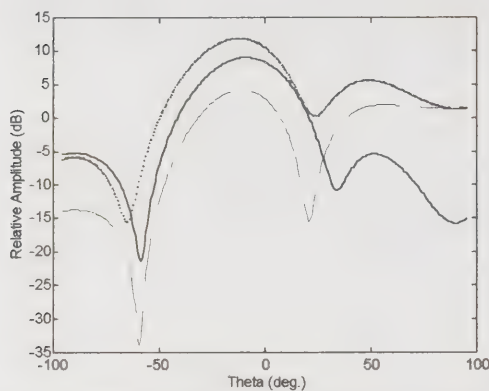


Fig. 6: Adapted response for SNR=20dB, with the null steered to direction ($\phi = -10^\circ$, $\theta = -60^\circ$), varying number of iterations. Solid curve=20, dashed curve=30 and point curve=10 iterations.

Conclusion

In this article, we have proposed a beamformer in which the weight vector is chosen so as to minimize output variance of beamformer output, subject to the constraint.

The adaptive process was developed for a planar antenna array through a neural linear based on the complex mean square algorithm. The LMS algorithm was used to weight adjustment in adaptive process.

Computer simulations for illustration and for examine the performance of the method have shown the success of the algorithm and it can be applied for a lot of antenna's structure in particular, the microstrip antennas because it always associated low cost, low weight, simple to manufacture and conformity to the supporting structure, but this type of antenna has

disadvantages as low radiation efficiency, small bandwidth and low gain. The adaptive process is going to be suitable for those antennas.

References

- [1] Po-Rong Chang, Wen-Hao Yang and Kuan-Kin Chan, "A neural network approach to MVDR beamforming problem," *IEEE Trans. Antennas Propagat.*, vol.40, pp. 313-322, March 1982.
- [2] Ching Tseng, "Minimum variance beamforming with phase-independent derivative constraints," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 285-294, March. 1992.
- [3] Zbynek Raida, "Steering an adaptive antenna array by the simplified Kalman filter," *IEEE Trans. Antennas Propagat.*, vol.43, pp. 627-630, Jan. 1995.
- [4] T. B. Vu, "On null steering in rectangular planar array," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 995-997, Aug. 1992.
- [5] Hao Ye and Ronald D. De Groat, "A generalized sidelobe canceller with Soft Constraints," *IEEE Trans. Signal Processing.*, vol. 40, pp. 2112-2116, Aug. 1992.
- [6] O. L. Frost, "Na algorithm for linearly constrained adaptive array processing," *Proc. IEEE, Trans. Antennas Propagat.*, vol. 60, pp. 926-935, Aug. 1972.
- [7] L. J. Griffiths and C. W. Jim, "An Alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp 27-34, Jan. 1982.
- [8] Ching-Yih Tseng and L. J. Griffiths, "A Unified approach to design of a linear constraints in minimum variance adaptive beamformers," *IEEE Trans. Antennas Propagat.*, vol.40, pp 1533-1542, Dec. 1992.
- [9] Shiann Yu and Ju-Hong Lee, "Adaptive Array beamforming based on a efficient technique," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1094--1101, Aug. 1996.
- [10] Ho Yang and Mary Ann Ingram, "Design of partially Adaptive arrays using the singular-value decomposition," *IEEE Trans. Antennas Propagat.*, vol.45, pp 843-850, May. 1997.
- [11] Simon Haykin, "Adaptive Filter theory," Prentice Hall, Third Edition, 1996.
- [12] A. M. Haimovich, and Yeheskel Bar-Ness, "An Eigenanalysis interference canceler," *IEEE Trans. Signal Processing.*, vol. 39, pp 76- 84 Jan. 1991.
- [13] Strang G. and Borre K. "Linear Algebra Geodesy and GPS," Wellesley-Cambridge Press, 1997.

A Dual-Polarized Microstrip Antenna Array with High Isolation Fed by Coplanar Network

ShiChang Gao* and ShunShi Zhong

Department of Communications Engineering, Shanghai University
20 ChengZhong RD, Jiading, Shanghai, 201800, P.R. China

Abstract In this paper, serial corner feeding of square patch with two ports is proposed to realize dual polarization. For the four-element dual-polarized array, a novel coplanar feedline network is presented. An efficient method of analysis is developed and a practical array is manufactured. The measured isolation is less than -40 dB at 6.07GHz. The proposed array occupies smaller area than other dual-polarized arrays commonly used and is easier to be further combined to form larger coplanar array.

1. INTRODUCTION

Several applications in communication and radar systems often require antennas with dual polarization capability. In satellite communication systems, frequency reuse involving the use of orthogonal polarizations gives enhanced capacity. In land mobile communications, the multipath propagation phenomenon leads to the deterioration of system performance, while polarization diversity provides an important way to combat the multipath fading. In radar, dual polarization gives improved clutter suppression. In all these cases, dual-polarized antennas are quite necessary.

There is increasing interest in microstrip antennas due to their many advantages such as compactness, low profile, mass production, light weight, conformity to surface, and finally, the possibility to integrate active components. As saving space, cost and weight is usually a premium, single layer structure is desirable. Serial corner feeding of printed array is proposed in [1] to reduce the size of feedline network. For dual-polarization operation, the common practice is to locate the two feeding ports at the central points of orthogonal edges of square patch, as it is assumed to achieve the highest isolation. Most dual-polarized microstrip antenna arrays are built according to this assumption, and the isolation of -20dB is achieved[2-3]. Many other dual-polarized arrays are also reported, where usually very complex structures of the antennas or feedline networks are used to realize high isolation[4-5].

In this paper, serial corner feeding of square patch with two ports is proposed to realize the dual-polarized microstrip antenna. A novel coplanar feedline network is also presented for the dual-polarized array. An efficient method of analysis is developed, which is based on the multiport network theory, the improved cavity theory and the segmentation technique[6]. A 2×2 microstrip antenna array is designed. The measured isolation is less than -40 dB at 6.07GHz. The advantages of the array include high isolation, small area, coplanar structure and easiness to be further combined to form larger array.

2.DUAL-POLARIZED MICROSTRIP ANTENNA ARRAY

2.1 Dual-Polarized Antenna

The geometry of the dual-polarized antenna is shown in Figure 1(a). It consists of a square patch with two ports located at two corners. The serial corner feeding method proposed in [1] is extended here to realize dual-polarization operation. Using this feeding method, it is easy to form large arrays with small feedline network. Let's assume that the substrate is electrically thin ($d \ll \lambda_0$). To analyze the antenna, the two-dimensional planar circuit approach is used, combined with the improved cavity theory and the segmentation technique [6]. The antenna is divided into five segments, as shown in Figure 1(b). Segment s1 and s2 are one microstrip line with dimension $s \times w$ and one right-angled isosceles triangle. Segment s3 is a square patch with dimension $a \times a$ respectively. The connection between two segments is discretized by interconnections only at a finite number of ports. The individual segments are treated as multiport planar networks. The Z-matrix for them are evaluated using their Green's functions. The Green's functions are given in [6]. Several modifications are made here. For segment s3, effective wavenumber k_{eff} and a_{eff} are introduced to replace k and a respectively, as presented in [6]. The ports of the individual segments are separated into external ports and connected ports. At the interface between two segments, the connected ports are equally divided into two groups, corresponding to the ports of two segments respectively. For symmetry, the Z-matrices of s4 and s5 are same as those of s2 and s1 respectively. e and f indicate the two input ports of the antenna. By the multiport connection method, the five segments are combined one by one [6]. We finally obtain the impedance matrix of the antenna as

$$\begin{bmatrix} V_e \\ V_f \end{bmatrix} = \begin{bmatrix} Z_{ee} & Z_{ef} \\ Z_{fe} & Z_{ff} \end{bmatrix} \begin{bmatrix} I_e \\ I_f \end{bmatrix}$$

where V_e, V_f and I_e, I_f are port voltages and port currents respectively.

2.2 Dual-Polarized Array

The dual-polarized array is presented in Figure 2. The 180° phase difference of each port is used. To study the array structure, segmentation technique is again applied. The array is divided into the segments α, β and γ , where the segment α denotes the four two-port patches, while the segments β and γ denote the feedline networks of the array for each polarization respectively. The ports of these segments are labelled differently. Based on the previous theory, the impedance matrix of the segment α can be easily computed. The segments β and γ are further divided into several smaller regular segments with known Green's functions, and their impedance matrices are computed by the segmentation technique. By combination of the three segments one by one, the Z-matrix of the whole array is derived as

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

and the isolation between the two input ports is obtained as

$$S_{21} = \frac{2Z_{21}Z_0}{(Z_{11} + Z_0)(Z_{22} + Z_0) - Z_{21}Z_{12}}$$

where Z_0 is the characteristic impedance of the feedline.

3.EXPERIMENTAL RESULTS

Based on the theory, a 2×2 array at 6 GHz is designed and manufactured. The design parameters are: $a = 15.0\text{mm}$, $d = 0.8\text{mm}$, $w = 0.4\text{mm}$, $\epsilon_r = 2.8$. The distance between the square patches is $0.76\lambda_0$. The antenna array is measured by HP8510-B Network Analyzer. Figure 3 shows the results of VSWR at two ports and Figure 4 shows the isolation between two ports. It is shown that isolation of -41.434dB is achieved at 6.07GHz. The theoretical results are also shown. The slight differences between them may be attributed to the manufactural discrepancies and the simplification of the theoretical model.

4 CONCLUSIONS

In this paper, a new dual-polarized array is presented. Serial corner feeding with two ports is proposed to realize dual-polarized antenna. A novel coplanar feedline network is also presented, together with a method of analysis for the array. The measured results demonstrate high isolation between the two input ports. The array has very simple structure and small, coplanar feedline network. It is easy to be further combined to form larger coplanar array.

REFERENCES

- 1 J.P Daniel, E. Penard, M. Nedelec and J.P Mutzig, Design of a low cost printed antenna array, Proc. ISAP'85, Kyoto, Aug, 1985, pp. 121-124
- 2 J.R James, P S Hall, C Wood, Microstrip antenna theory and design, Peter Peregrinus Ltd, 1981, pp. 166-167
- 3 L.J du Tolt and J.H Closte, Dual polarized linear microstrip patch array, IEEE Antennas Propagat. Soc, Int. Symp. Dig., Vol. 2, June, 1987, pp. 810-813
- 4 P. Brachet and J.M. Baracco, Dual-polarization slot-coupled printed antennas fed by stripline, IEEE Trans. Vol. AP-43, July, 1995, pp. 738-742
- 5 P. Piazzesi, S. Maci and G. Gentili, Dual-band dual-polarized patch antennas, Int. J. MMCAE, Vol. 5, No. 6, 1995, pp. 375-383
- 6 K.C. Gupta, Multiport network modeling approach for microstrip patches and arrays, in Handbook of microstrip antennas, J.R. James and P.S. Hall ed., Peter Peregrinus, 1989

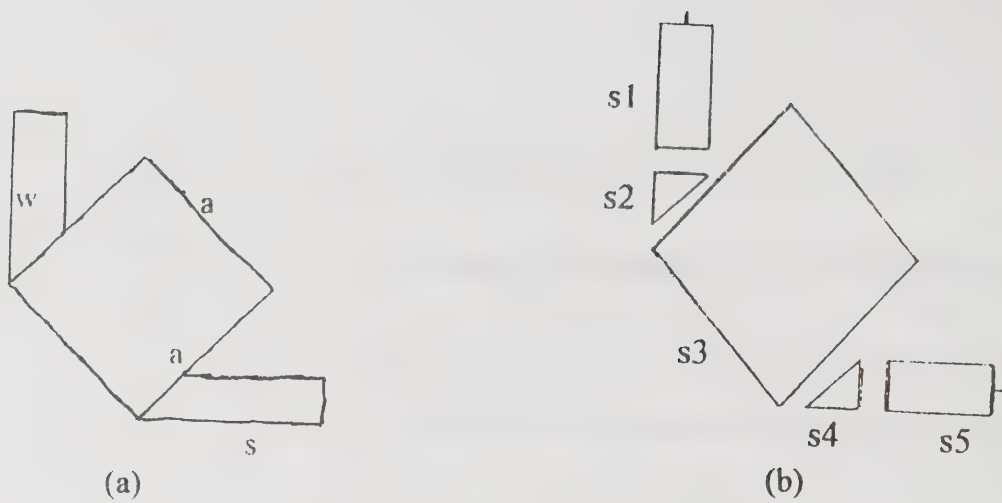


Figure 1 Dual-polarized antenna

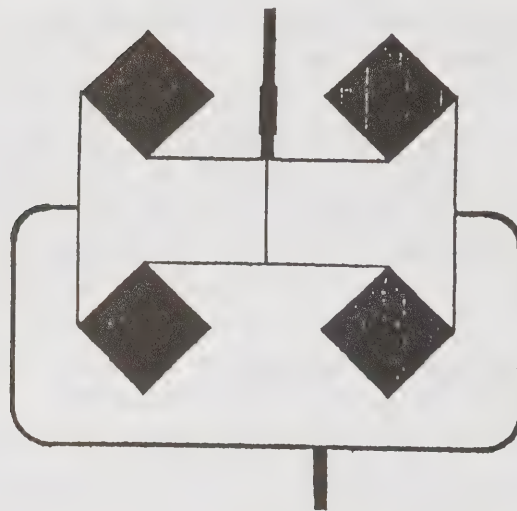


Figure 2 A 2×2 microstrip antenna array for dual polarization

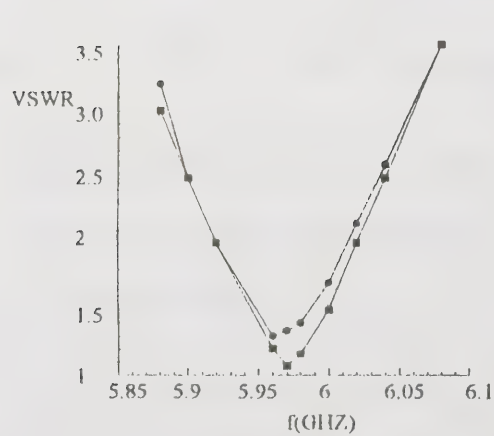


Figure 3 VSWR of two ports
port 1
port 2

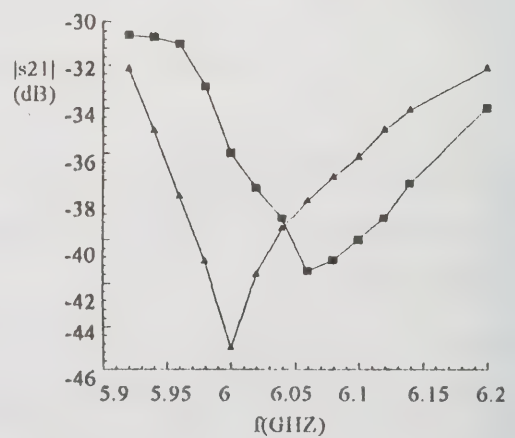


Figure 4 Isolation between two ports
measured
theory

Low-cost, Dual-band, Handset Antennas for LEOS Communications

James McLean, Gentry Crook
Tactical Systems Research, Inc.
3207 Yellowpine
Austin, TX 78757

tel: (512) 450-0544, fax: (512) 450-0415, e-mail: mclean@ccsi.com

Siegfried Mikuteit
Torrey Science Corporation
3550 General Atomics Court, Bldg. 14
San Diego, CA 92121

Abstract

A new, low-cost, mass-produceable, dual-band antenna design has been developed. One embodiment of the design provides passbands centered at 305 MHz (295-315 MHz) and 380 MHz (370-390 MHz) for use with the MUBLCOM (Multiple-path Beyond Line-of-sight Communications) LEOS (Low Earth Orbiting Satellite) communications system. However, because the new design provides improved flexibility in the spacings of the two passbands, it may be used to provide passbands with smaller or larger spacings. Furthermore, the antenna is suitable for either printed circuit or coaxial realization, thus allowing for low-cost mass production.

Keywords: Dual-band antenna, Mobile antenna, MUBLCOM, LEOS communications, Handset, Handheld

Introduction

Dual-band antennas are required for many different types of mobile communication systems including (but not limited to) some Low Earth Orbiting Satellite (LEOS) systems such as the Multiple-path Beyond Line-of-Sight Communications (MUBLCOM) system. When the two bands are widely spaced or are nearly harmonic, implementation of a dual band antenna is straightforward. However, moderately-spaced bands at VHF and lower UHF frequencies present a problem for mobile antenna design. For example, the MUBLCOM system requires two bands, 295-315 MHz and 370-390 MHz. Each band has approximately 5-6% fractional bandwidth and the ratio of the center frequencies is approximately 1.25:1. Furthermore, because of the unique features of the MUBLCOM system, both bands can be used for transmit from handheld units and thus good input match is essential over both bands. This is in contrast to systems such as AMPS where one band is used solely for mobile reception and thus the requirements for antenna input match are relaxed for that band. Because of fundamental bandwidth/size constraints [1] it is usually not possible to design an antenna with a contiguous passband large

enough to cover both bands which is still small enough for use with a mobile or handheld unit. In this case, contiguous coverage of both bands would require nearly 28% fractional bandwidth. Thus it is necessary to utilize a dual-band design which exhibits rejection in the region between the passbands in order to conserve so-called "matching area" [2].

Background

One effective means of constructing a dual band antenna is by placing a parallel LC resonator in series with a resonant radiating element (such as a dipole or monopole). The drawbacks to this approach are (1) complexity and (2) low electrical efficiency due primarily to losses in the inductor. The new design presented here is instead an extension of the conventional open, folded wire monopole [3]. An extensive discussion of conventional folded antennas is given in [4]; here only salient details are presented in the interest of brevity.

In a folded antenna, the total current on the structure can be thought of as consisting of a radiating mode and a transmission line mode [4]. In the more common form of folded monopole or dipole (such as that often used for FM

broadcast reception), the undriven element(s) are shorted to ground. The transmission line and transformed radiating mode impedances therefore appear in parallel at the input, and proper alignment of the resonance frequencies of these modes can provide a relatively broad single passband. The open, folded monopole, on the other hand, effectively places the transmission line and radiation mode impedances in *series* at the input terminals. That is,

$$Z_{in} = Z_{rad} + \frac{1}{4}Z_{tl} \quad (1)$$

where Z_{rad} is the radiation mode impedance and Z_{tl} is the transmission line mode impedance. The transmission line mode impedance is that of a short-circuited transmission line

$$Z_{tl} = Z_0 \tan \beta l \quad (2)$$

where Z_0 is the characteristic impedance of the transmission line, β is the wavenumber, and l is the length. Thus the transmission line mode is open-circuited at the frequency at which it is one-quarter wavelength long. Below the quarter-wavelength frequency, the transmission line mode impedance is inductive while above (but still below the half-wavelength frequency) it is capacitive. In this case, proper alignment of the transmission line and radiation mode resonant frequencies will result in the formation of two distinct passbands.

This design allows significant flexibility in adjusting the positions of the two passbands, in addition to providing size reduction. The characteristic impedance of the transmission line mode is the primary determining factor of the spacing between the two passbands, and dielectric loading provides the ability to vary this characteristic impedance over a large range. By utilizing a relatively low characteristic impedance, the passband spacing may be narrowed. Furthermore, the dielectric loading reduces the physical length of the folded section required for resonance. The resonance frequency of the radiation mode may be lowered by any of several means, including a series tuning element and/or the addition of a capacity (top) hat. Thus a reduced-size dual-band antenna can be provided.

In the current design, which is shown in Figure 1, a resonant monopole (essentially, one-quarter wavelength long at the geometric mean of the two passbands, 340 MHz) is combined with a low-impedance sleeve resonator ($Z_0 \approx 5\Omega$, quarter-wave frequency: 340 MHz). Because the radiation resistance is relatively high (near 100Ω) and increasing with frequency, a low-pass impedance trans-

forming network is used. Because it was desired to maintain the cylindrical design of the antenna, the tuning network was realized with coaxial elements. Thus, the low-pass impedance transformer network consists of a series coaxial inductor and a shunt coaxial capacitor. An equivalent circuit for the antenna is shown in Figure 2. This design can also be implemented in printed circuit form. In this case, the quarter-wave resonator sleeve can be realized by two conductors on opposite sides of a printed circuit board. Prototypes were fabricated using standard 31 and 62 mil FR-4 board which easily accommodated the low transmission line mode impedances which were required to provide the closely-spaced passbands.

Dipole designs which utilized a choke sleeve to provide isolation from and hence operation essentially independent of the radio housing were also developed. Such designs are particularly desirable because they greatly reduce the magnitude of RF current induced on the radio housing and thus limit user specific absorption rate (SAR) of RF energy. However, these designs were deemed too large for use with handheld units.

Experimental Results and Conclusions

Several prototype antennas have been fabricated and experimentally characterized. In Figure 3 the input return loss of one prototype is given, showing the two passbands. In Figure 4 the predicted directional gain of this prototype is presented for three different environments: (1) antenna mounted on the radio in free space, (2) antenna mounted on radio 2 meters over a perfectly conducting ground, (3) antenna mounted on radio 2 meters over lossy earth. Simulations and experiment indicate that the directivity of the antenna is strongly affected by the mounting position on the radio as well as the presence of the user. However, as expected, the input match is only modestly affected by these factors. As can be seen in Figure 4, the gain on the horizon is strongly affected by the near cancellation of the the line-of-sight ray with the ground-bounce ray. This is significant as the LEOS satellites are often at low elevations.

In summary, a dual-band antenna for handheld LEOS communications has been developed and successfully demonstrated. The input return loss is better than 13 dB over each passband. Electrical efficiency and gain also quite good; maximum gain is better than 2 dBi. Further improvements may be obtained by combining several such antennas to provide antenna diversity. This would combat both fading caused by multipath interference as well as shadowing caused by the user.

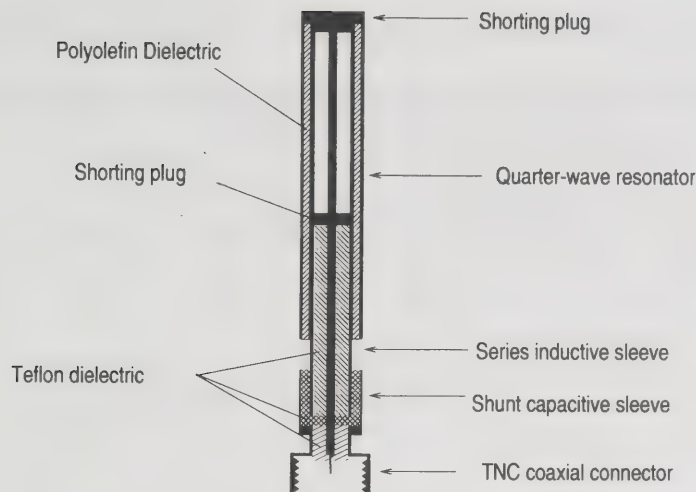


Figure 1: Dualband monopole design

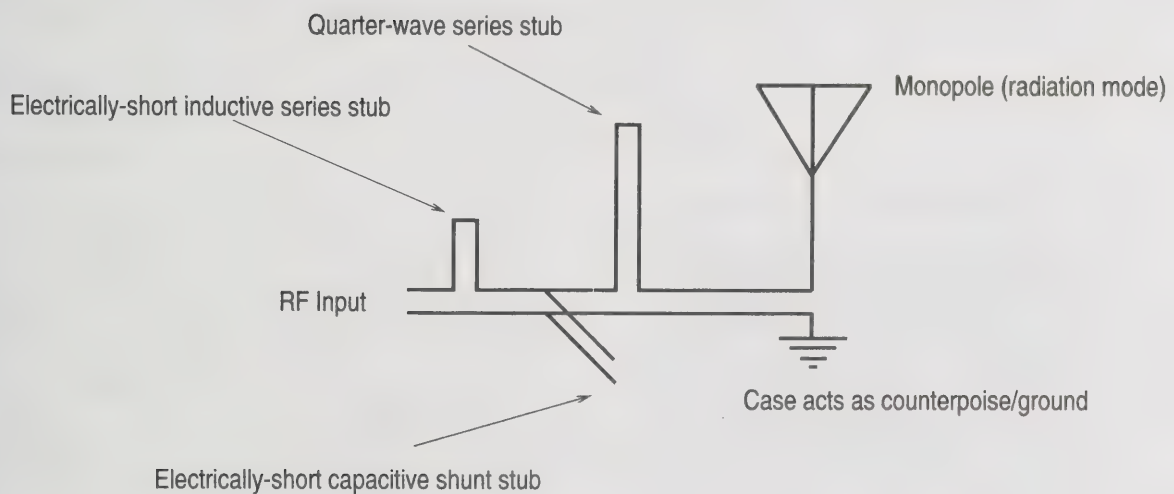


Figure 2: Equivalent circuit of dual-band monopole

References

- [1] J. S. McLean, "A Re-examination of the Fundamental Limits on the Radiation Q of Electrically-small Antennas," *IEEE Trans. Ant. Prop.*, Vol. 44, No. 5, pp. 672-677, May 1996.
- [2] H. A. Wheeler "The Wide-band Matching Area for a Small Antenna," *IEEE Trans. Ant. Prop.*, Vol. AP-31, No. 2, pp. 364-367 March 1983.
- [3] Chen To Tai and Stuart A. Long, "Dipoles and Monopoles," in *Antenna Engineering Handbook*, Richard C. Johnson, Ed., McGraw-Hill, Inc. 1993.
- [4] C. A. Balanis, *Antenna Theory, Analysis and Design*, John Wiley & Sons, New York, 1982, pp. 439-444.

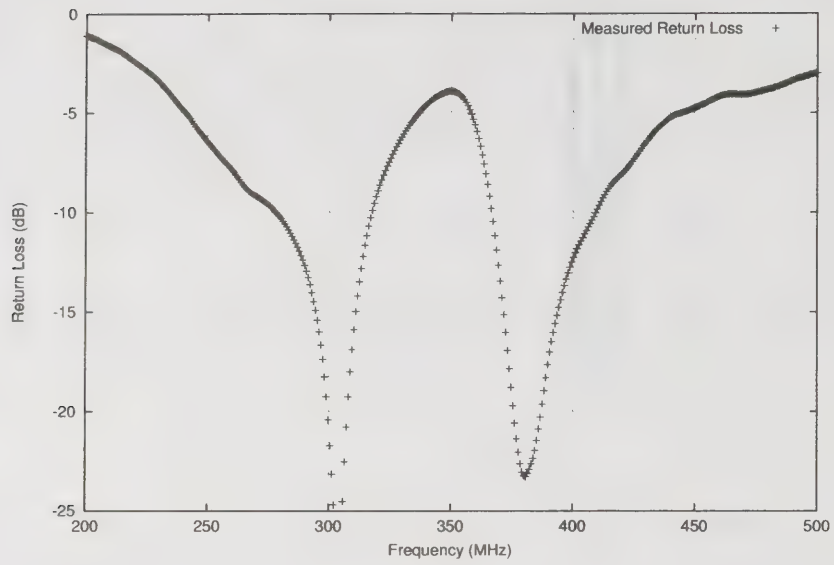


Figure 3: Measured Return Loss

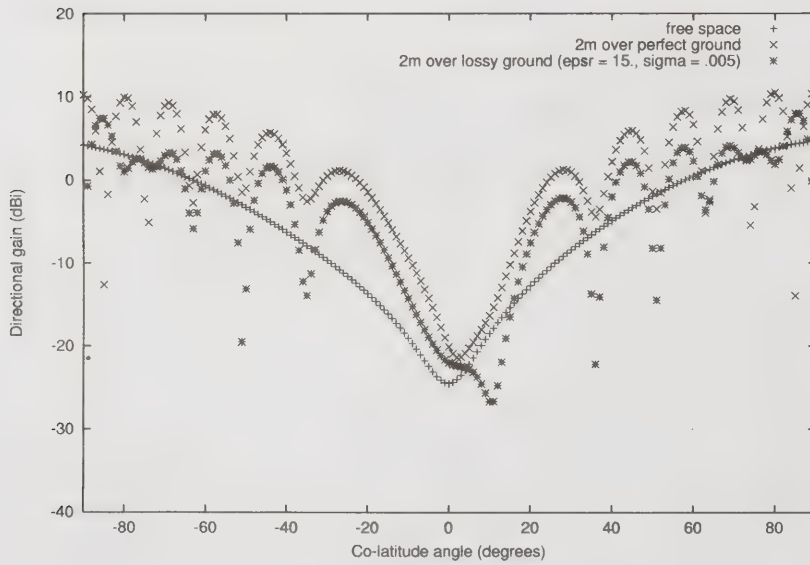


Figure 4: Directional Gain (Elevation cut)

A Novel Approach for Gain and Bandwidth Enhancement of Patch Antennas

Yongxi Qian, Dan Sievenpiper, Vesna Radisic, Eli Yablonovitch and Tatsuo Itoh
Department of Electrical Engineering
University of California, Los Angeles
405 Hilgard Avenue, Los Angeles, CA 90095
Phone: (310)206-5670, Fax: (310)206-4819, Email: yqian@ucla.edu

Abstract

The microstrip patch is one of the most preferred antenna structures for low-cost and compact design of wireless communication systems and RF sensors. To overcome several intrinsic limitations of the patch antenna such as narrow bandwidth, low gain, and degradation of radiation efficiency at higher frequencies, we propose a novel technique for gain and bandwidth enhancement based on the photonic band-gap (PBG) concept. The Ku-band prototype demonstrates over 3 times bandwidth improvement, and 1.6 dB higher gain or 45 % increase in effective radiated power (ERP), compared with a regular patch with identical dimensions. System design issues such as co-site interference can also be alleviated by the improved beam patterns of the new PBG antenna.

Key Words: Patch Antenna, Gain, Surface Wave, Radiation Efficiency, Photonic Band-Gap (PBG)

Introduction

Microstrip patch antennas offer an attractive solution to compact, conformal and low-cost design of many wireless application systems such as ATM Wireless Access (AWA) [1] and millimeter-wave automobile sensors [2]. For easy integration with the RF front-end, patch antennas on substrate with high dielectric constant (Si, GaAs, etc.) are preferred. However, this usually results in very narrow bandwidth (~1%), and degradation in radiation efficiency and total antenna gain, particularly at higher frequencies. Recently there has been a great interest in realizing high efficiency patch antennas on high permittivity substrates, including the use of the latest micromachining technology [3].

We propose here a novel technique for gain and bandwidth enhancement of patch antennas based on the concept of photonic band-gap (PBG) structures [4]. By surrounding the patch antenna with a square-lattice of small metal pads with grounding vias, we observe a substantial suppression of surface waves excited in the dielectric substrate, which not only improves the antenna gain or effective radiated power (ERP), but also increases significantly its frequency bandwidth. Reduction in mutual coupling and co-site interference are some other potential benefits of the newly proposed PBG antenna.

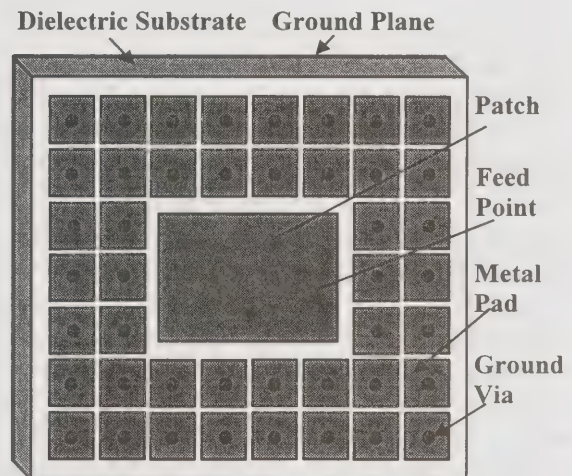


Fig. 1 Schematic of the proposed microstrip patch antenna surrounded by a PBG lattice.

Design of the PBG Patch Antenna

Fig. 1 shows the schematic of the proposed PBG patch antenna surrounded by a square lattice of metal pads with grounding vias. A distinctive stopband exists for frequencies above the resonance of the two-dimensional PBG lattice, similar to that of the 3D structure as described in [4]. As a result, surface waves excited in the dielectric substrate will be suppressed, and improvement in both radiation efficiency and beam patterns can be expected.

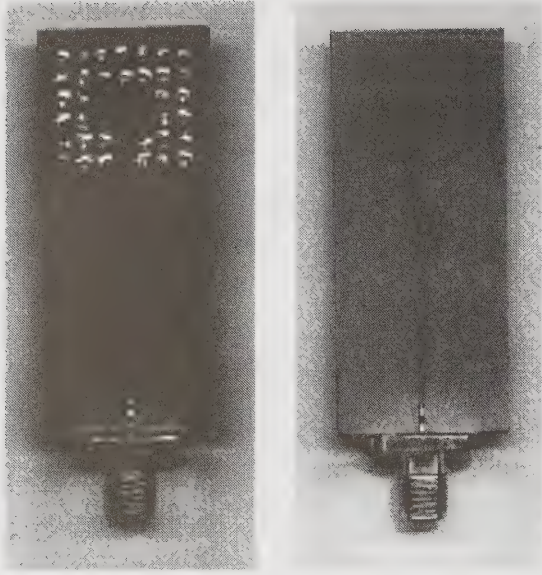


Fig. 2 Ku-band prototype of (a) the proposed PBG patch and (b) a reference patch antenna.

To confirm the proposed design concept, a Ku-band prototype was fabricated and tested. The antenna design started from a reference patch antenna on 25 mil thick Duroid with $\epsilon_r=10.2$. An inset feed scheme is employed to match the patch antenna to a $50\ \Omega$ microstrip feedline. For a center frequency of 14 GHz, the size of the patch antenna was found to be 120 mil long and 168 mil wide by using an FDTD simulation code [5]. Meanwhile, the PBG patch has the same dimensions with the reference antenna, except for the surrounding PBG lattice. The PBG pads are 88 mil squares with 8 mil gaps, and the grounding via in the center of each pad has a diameter of 8 mil. Fig. 2 shows the pictures of the fabricated PBG and reference patch antenna. The total width of the antenna substrate is 800 mil, corresponding to approximately one free-space wavelength at the center frequency. A relatively long (1.4 in) microstrip feedline is used to facilitate the E-plane pattern measurement.

Fig. 3 shows the measured input return loss of the two antennas. The reference patch has a minimum return loss of -11.6 dB at 14.0 GHz, and bandwidth (VSWR<2) of 1.6 %. Although the return loss level can be improved by further optimization of the inset feed position, the bandwidth will remain narrow (around 1~2 %) on

this type of high permittivity substrate, as also demonstrated recently in [3]. On the other hand, the PBG patch antenna measured a peak return loss of -24.4 dB at 14.75 GHz, and a bandwidth of 5.4 %, which is 3.4 times wider than that of the reference patch.

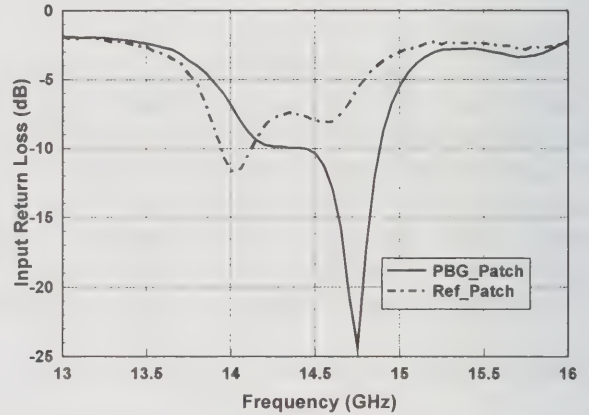


Fig. 3 Measured input return loss of the PBG and reference patch antennas.

Radiation Patterns

Fig. 4 shows the measured the H- and E-plane radiation patterns of the two patch antennas, including both co- and cross-polarization patterns. While the patterns have been normalized here for easy visualization, the peak power received by the PBG patch is 1.8 dB higher in the H-plane and 1.6 dB higher in the E-plane. The measurement was taken at 14.15 GHz where the two patches have identical return loss (-9.4 dB). Since all other parameters are the same, the increased power indicates a gain enhancement with the PBG antenna, which will be discussed further in the next section.

As can be seen from Fig. 4, the PBG patch has reduced radiation power along the dielectric substrate (90 degree from broadside) and smaller ripples in its E-plane pattern, indicating an effective suppression of surface waves. The radiation power at the H-plane edges remains the same as the reference patch since no surface wave is excited and propagates in that direction. The backlobe of the PBG patch is lower than that of the reference because of the same surface wave suppression effect. The slight asymmetry in the E-

plane patterns are due to the blocking of the SMA connector. Meanwhile, the PBG antenna does show slightly higher cross-polarization compared to the reference patch. Nevertheless, the -19.1 dB (H-Plane) and -14.7 dB (E-Plane) cross-polarization levels are still reasonably low, and may have been caused partly by the uneven surface of the antenna due to the soldering of the grounding vias, which can be improved by using alternative fabrication techniques such as plating.

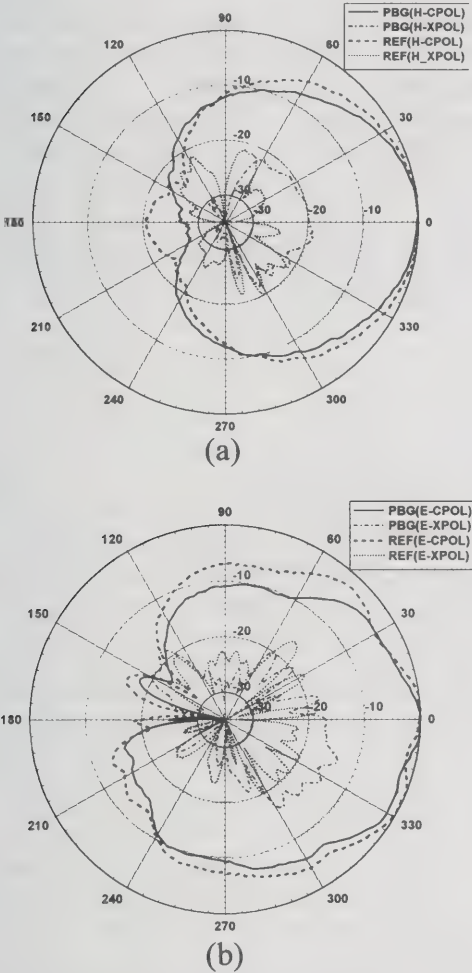


Fig. 4 Measured radiation patterns at 14.15 GHz for the PBG and reference patch antennas: (a) H-Plane and (b) E-Plane.

Gain and Radiation Efficiency

The gains of the proposed PBG and reference patch antennas are measured by using the gain-transfer technique [6]. Based on the relative gain of the patch measured with respect of a standard-gain horn antenna, we can de-embed

the net gains of the two patch antennas, as shown in Table I. The de-embedded gains of the PBG and reference patch are 6.77 dB and 5.16 dB, respectively, showing a 1.61 dB gain enhancement with the PBG antenna. This is equivalent to a 45 % increase in the effective radiated power (ERP) by the PBG antenna, assuming that everything else is the same in the whole RF system. This measurement result is also in consistence with the pattern measurement results described in the previous section.

Meanwhile, by integrating the radiation patterns as shown in Fig. 4 to obtain the antenna directivities, and comparing them with the measured gains, we can calculate the radiation efficiencies of the two antennas. The calibrated radiation efficiencies of the PBG and reference antennas are found to be 85 % and 83 %, respectively. It should be pointed out here that the radiation efficiency measured in this way is the efficiency of total radiation, which should be close for both patch antennas, since they are expected to dissipate similar amount of metal and dielectric losses. The surface wave suppression effect as evidenced by the radiation patterns shown in Fig. 4, however, is important because it reduces the radiation energy along the substrate and backside of the antenna, which is the major cause of mutual coupling in antenna arrays, or co-site interference within multiple RF-system environments.

It should also be pointed out that although it is possible to realize higher gain using an array of multiple patches, the associated feeding network for such an array is not only complicated but will also introduce additional feedline losses which tend to reduce the gain. Also, the array approach does not provide a direct solution to the surface wave problem as mentioned above.

Conclusion

In conclusion, we have proposed and demonstrated a novel approach for gain and bandwidth enhancement of microstrip patch antennas. The design is compact and robust, requiring only one single layer of dielectric material, and is compatible with standard planar fabrication technology. The multi-fold improvement in antenna performances (wider bandwidth, improved gain, lower backside

radiation, beam shape control and surface wave suppression) resulted from this single approach indicates great potentials, especially for applications at microwave and millimeter-wave frequencies.

Acknowledgment

The authors thank Prof. Yahya Rahmat-Samii for kindly providing his antenna range. This work was supported by US Army Research Office MURI under contract DAAH04-96-1-0389 and DAAH04-96-1-0005.

References

- [1] T. Seki, K. Uehara, M. Sato and Y. Konishi, "Technology of antennas for high speed wireless access," *MWE'97 Microwave Workshop Dig.*, pp. 181-186, Yokohama, Japan, Dec. 1997.
- [2] M. Singer, K. M. strohm, J. -F. Luy and E. M. Biebl, "Active SIMMWIC-antenna for automotive applications," *IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 1265-1268, Denver, CO, June 1997.
- [3] G. P. Gauthier, A. Courta y and G. M. Rebeiz, "Microstrip Antennas on Synthesized Low Dielectric-Constant Substrates," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1310-1314, Aug. 1997.
- [4] D. Sievenpiper and E. Yablonovitch, "Eliminating surface currents with metallodielectric photonic crystals," to be presented at *1998 IEEE MTT-S Int. Microwave Symp.*, Baltimore, MD, June 1998.
- [5] E. Yamashita and Y. Qian, *Analysis of Microwave Circuits and Planar Antennas Using the FDTD Method*, Realize Inc., Tokyo, May 1996.
- [6] C. A. Balanis, *Antenna Theory - Analysis and Design -*, John Wiley & Sons, New York, 1982.

Table I
Measured Gain of the PBG and Reference Patch Antenna

	PBG Patch	Reference Patch
Frequency (GHz)	14.15	14.15
Measured Gain (dB)	5.02	3.41
Mismatch Loss (dB)	0.53	0.53
Connector Loss (dB)	0.38	0.38
Feedline Loss (dB)	0.84	0.84
De-Embedded Gain (dB)	6.77	5.16

Performance of Lens Antennas in Wireless Indoor Millimeter Wave Applications

Carlos Fernandes[†] and José Fernandes[‡]

[†]Instituto de Telecomunicações - Pólo Lisboa, Av. Rovisco Pais, 1096 Lisboa Codex - Portugal
Tel.: +351 1 8418481 Fax: +351 1 8417284 E-mail: carlos.fernandes@lx.it.pt

[‡]Instituto de Telecomunicações - Pólo Aveiro, Universidade de Aveiro, 3810 Aveiro - Portugal
Tel.: +351 34 377900 Fax.: +351 34 377901 E-mail: zf@ua.pt

Abstract

Dielectric lens antennas can be designed to produce highly shaped beams that significantly improve the performance of emerging wireless indoor millimeter wave systems. A lens configuration is analyzed in this paper that produces a circular symmetric cell with uniform spatial distribution of power, with fairly sharp boundaries and scalable cell radius. The last characteristic is used to control the reflections at side walls. An hemispherical coverage lens antenna is designed for the mobile terminal to ensure relatively free movement. The impact of these antenna is analyzed in terms of cell coverage and channel time dispersion, considering the effect of cell radius scaling, and mobile terminal antenna tilting. Measurements and simulations show that the proposed lenses outperform common solutions based on pyramidal horns or biconics.

Key words: Lens antennas, shaped beams, millimeter waves, wireless applications

Introduction

The wide acceptance of mobile services by the users and strong competition in the field has led to an explosive growth in the cellular and mobile telephony in the last ten years. The increasing needs of larger bandwidths to support broadband services has been a major driving force pushing the development of Mobile/Wireless Broadband Systems, aiming to extend to the mobile users the wide range of services available in the Broadband Integrated Services Digital Network (B-ISDN). Due to the saturation of the lower part of the frequency spectrum this type of system will operate in the millimeter-wave band.

It has been shown that at these frequencies the antennas have a significant impact on the characteristics of the multipath radio channel and therefore they play a key role in the system performance[1]-[3]. The power available from solid state devices is limited so it should be fairly uniform distributed over the cell while keeping the channel time dispersion at low extent, since it directly impacts on the achievable carrier bit rate. Adaptive equalization is unavoidable in high bit rate systems, but it cannot cope with large time dispersion values often encountered in typical scenarios. Hence, some directivity is required both at the base and mobile station antennas to favor the link budget and to cooperate with the equalizer in the mitigation of the channel time dispersion. But the antenna directive requirements must not entail a restriction of terminal mobility.

In this paper we analyze the impact of using a novel and quite cheap antenna configuration for indoor scenarios that is able to cope with the above requirements. The basic idea is to use a dielectric lens antenna at the base station (BST) to produce a secant squared ($\sec^2 \theta$) type of radiation pattern within the cell limits in order to compensate the free space attenuation at each direction. The BST antenna is paired with an hemispherical pattern lens in the mobile terminal (MT), so that the average received power remains reasonably constant for all positions of the mobile or portable terminal within the cell.

The proposed lens combination further provides very sharp cell boundaries, with negligible radiation outside the cell limits. A remarkable characteristic of $\sec^2 \theta$ patterns is that cell dimensions are scaled to the antenna height. This provides a simple means to control the illumination of the walls at the edges of the cell to maintain an adequate compromise between multipath effects and the need for alternative paths in case of line of sight (LOS) blockage.

Lens Antenna

The basic principles for the design of an axis symmetric amplitude shaping dielectric lens are addressed in [4]. The lens is fed by the aperture of a circular metallic waveguide which is embedded in the lens body. The lens surface is conveniently shaped to transform the feed aperture radiation pattern into the desired output beam. This antenna configuration is quite flexible allowing the design for different target

patterns from secant squared to hemispherical type, with linear or circular polarization.

The lens antenna design involves a two step procedure based on Geometrical Optics and on Physical Optics. Specific software tools were developed which allow a reliable calculation of the lens profile and prediction of corresponding radiation characteristics.

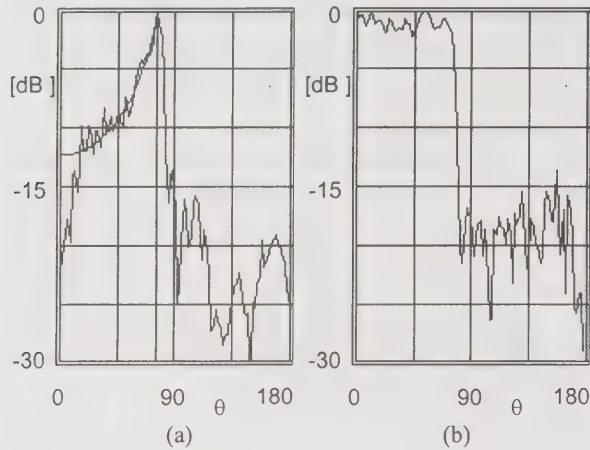


Figure 1 - Elevation power patterns of dielectric lens antennas measured at 62.5 GHz; lens diameter: 66 mm; $\epsilon_r=2.53$; a) BST antenna pattern superimposed on $\sec^2\theta$ target pattern. b) MT antenna pattern.

The BST dielectric lens is excited by the TM_{01} mode of the circular waveguide, producing a circular symmetric cell with vertically polarized field (dominant E_θ component). Figure 1.a shows the corresponding elevation pattern measured at 62.5 GHz, superimposed on $\sec^2\theta$ target pattern. The ground direction corresponds to $\theta \approx 0$ and the horizon to $\theta \approx 90^\circ$. The radiation is intentionally limited up to $\theta_M \sim 75^\circ$, which corresponds to 11 m diameter cell when the height difference between BST and MT antennas ($h_b - h_m$) is $\Delta h = 1.5$ m. The agreement between measured and target pattern is satisfactory except near $\theta \approx 0$, due to intrinsic limitation of linear polarization; prototypes using circular polarization do not show this problem. Notice the remarkable steep fall of radiation for $\theta > \theta_M$.

The MT lens is excited by the TE_{11} mode of the circular waveguide with circular polarization. The target pattern in this case is flat top up to $\theta_M \sim 75^\circ$, with no radiation outside this interval. The measured pattern shown in Figure 1.b complies with this specification. In this case $\theta \approx 0$ corresponds to the ceiling direction, and $\theta \approx 90^\circ$ to the horizon. Again notice the steep fall of radiation for $\theta > \theta_M$ which prevents ground pick up.

Cell coverage

CW measurements were performed in a 10.9 m x 8.8 m x 4.0 m room to evaluate the proposed antenna

configuration. Ceiling and side walls material is concrete, and the floor is covered with 1" thick ceramic tiles. An anechoic chamber with laminated wood external walls, tables, shelves and equipment racks are the main objects in this room.

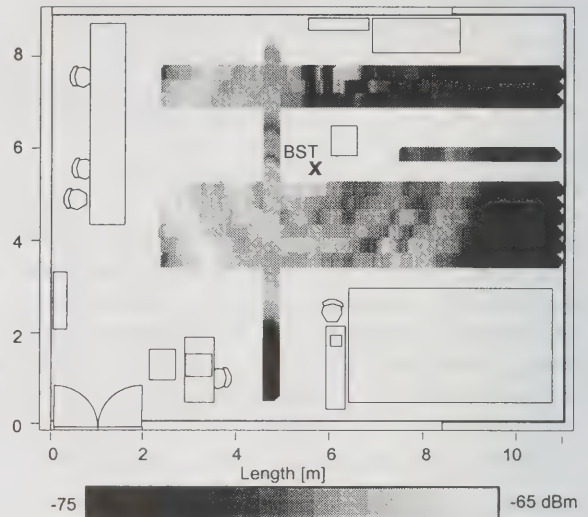


Figure 2 - Received power measured along selected paths in the test room, for $\Delta h = 1$ m, $h_b = 2.5$ m

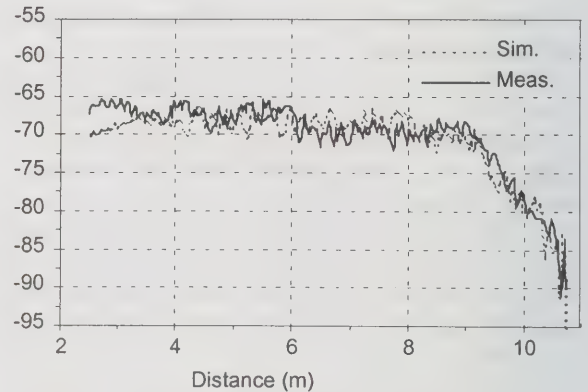


Figure 3 - Received power along a longitudinal path in the test room for $\Delta h = 1$ m, $h_b = 2.5$ m.

Figure 2 shows the BST antenna location and the measured received power distribution along linear paths taken in the uncluttered part of the room, for $\Delta h = 1$ m. The circular symmetric nature of the cell is quite apparent. The received power level is reasonably constant within about a 3.5 m radius, and falls off rapidly outside this region. A quantitative measure of this behavior is given in Figure 3, that corresponds to the closest path to BST taken along the room length. Simulation results obtained with a ray tracing tool described in [5] are superimposed, showing a good agreement. Despite the omnidirectional characteristic of BST and MT antennas, fading depth within the cell is negligible due to weak illumination of the walls.

Further simulations were performed in an empty but otherwise similar room for different BST antenna heights (h_b) and MT antenna tilting angles (γ_m) to demonstrate the performance of this antenna configuration as well as the allowed user movement freedom. The simulation uses measured antenna radiation pattern data. The MT antenna height (h_m) is always kept at 1.5m. Figure 4 shows the distribution of the normalized received power (NRP) calculated for $h_b = 3$ m. The cell is larger than in Figure 2 as expected because in this case Δh is increased by a factor of 1.5, but previous conclusions about the uniformity of power distribution and sharp cell boundary apply again.

The corresponding CDF curve in Figure 5 reveals that in 80% of the room area NRP > -69 dB and only 10% is below -73 dB. For $h_b > 3.5$ m, the NRP distribution tends to be uniform all over the room since the circle defined by the BST antenna radiation pattern covers all room area. However, as h_b increases the channel time dispersion (here represented by the sliding delay window parameter, SDW 90%) also increases as the MT picks up more and more significant multipath components mainly reflected from the side walls (Figure 5). This is an important characteristic of the proposed antenna configuration since alternative paths can be provided in a controlled way through h_b to cope with possible LOS blockage (Figure 6).

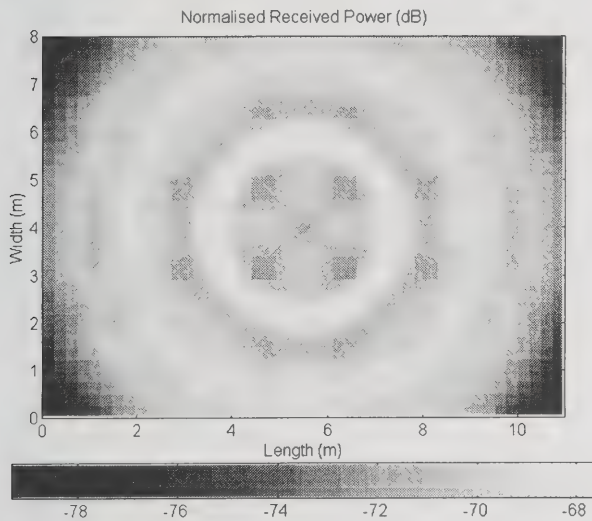


Figure 4 - Simulated NRP distribution along the room for $\Delta h = 1.5$ m, $h_b = 3$ m.

Figure 7 depicts the CDF of the NRP and SDW obtained along the room for several γ_m values, with $h_b = 3.5$ m. The NRP degradation is not significant for $\gamma_m < 5^\circ$, but it worsens for higher angles in about 50% of the room area, where the MT antenna main lobe fails to contain the direction of the BST antenna. The degradation is not very significant in the case of non LOS since the LOS component was already not

available: 2 dB at 10% probability for the NRP and 10 ns at 90% probability for the SDW.

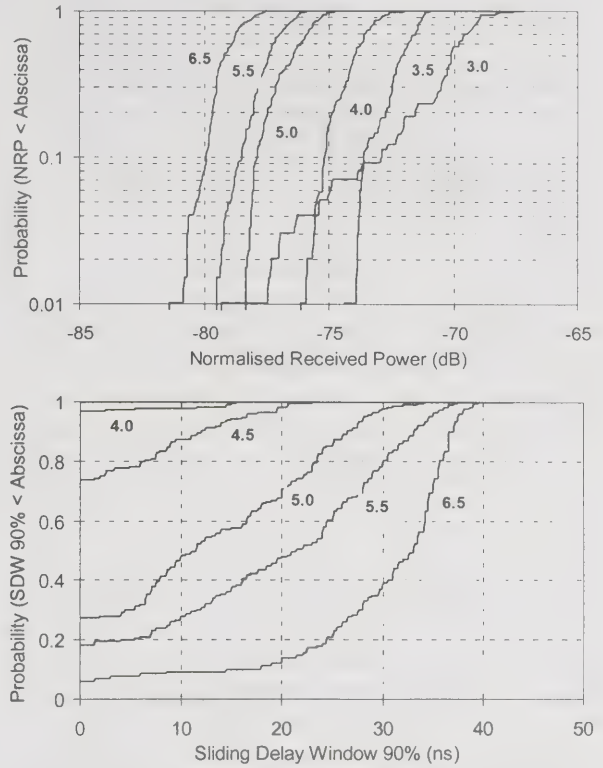


Figure 5 - CDF of the NRP and SDW parameters for several h_b values, in meters.

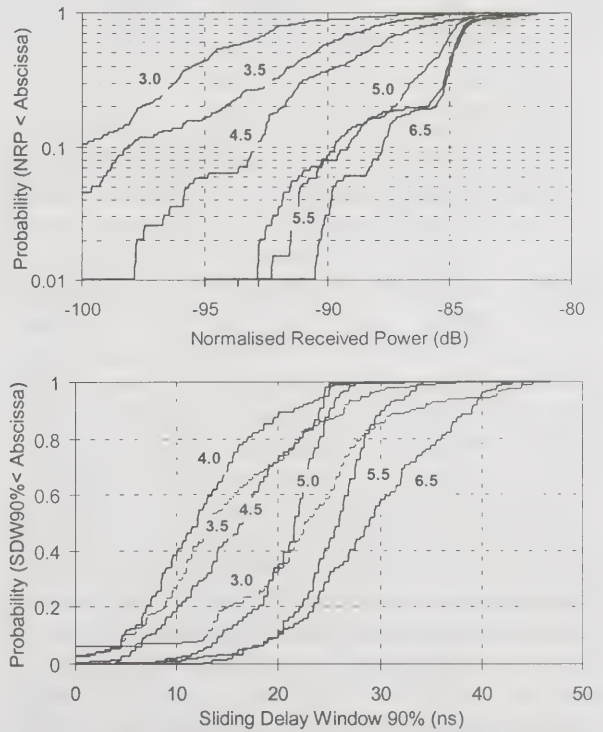


Figure 6 - CDF of the NRP and SDW parameters for several h_b values, in meters. LOS path is suppressed.

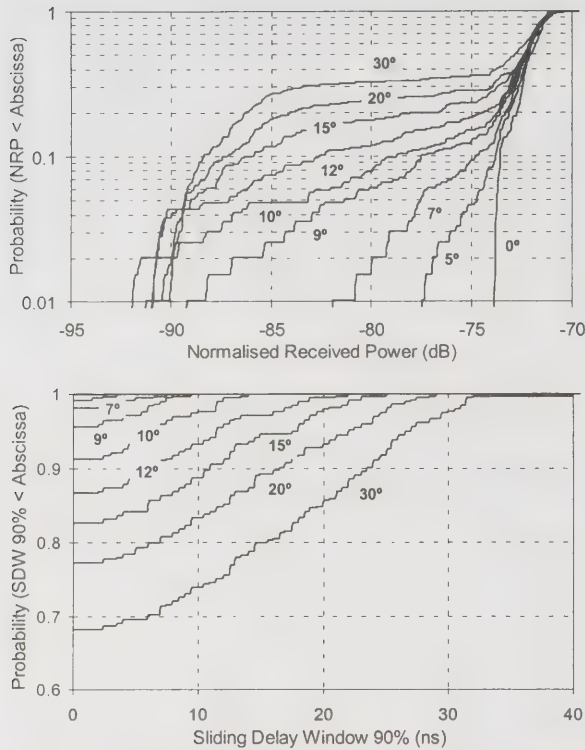


Figure 7 - CDF of the NRP and SDW parameters for several tilting angles γ_m of the MT antenna; $\Delta h = 2$ m, $h_b = 3.5$ m.

The tilting freedom of the MT was also studied for $h_b = 5$ m, showing a relatively small degradation due to the presence of several multipath components: 1 dB at 10% probability for the NRP and 3 ns at 90% probability for the SDW for γ_m up to 20° .

Impact on system performance

Considering a system with transceiver specifications used in [1-2], i.e., a decision feedback equalizer (DFE) with 7 taps in the forward and feedback filters with a $T_s/2$ (T_s : symbol time) tap spacing in the receiver, and using the method described in [1] to estimate the maximum carrier bit rate (CBR), one concludes that a minimum CBR of 170 Mbit/s (SDW < 40 ns) can be achieved in this room with the proposed antennas, even in non LOS condition, provided that the adequate h_b value is chosen. In LOS situation and with low MT antenna tilting the CBR can be much higher since the channel time dispersion is near zero for lower antenna heights. On the other hand, a significant high bit rate can be achieved without equalization: using the rule of thumb that the rms delay spread (DS) cannot exceed 10% of T_s [6], a symbol rate of 10 Msymbols/s can be achieved in LOS condition with γ_m up to 12 degrees for $h_b = 3.5$ m. For the sake of comparison, taking the values of DS (~ 32 ns) and DW (~ 100 ns) parameters obtained with biconic horn antennas at the BST and at the MT in a room with similar dimensions (11 m long by 7 m

width, see [5]) one concludes that for biconics the maximum CBR results in about 70 Mbit/s using the equalizer and 3 Msymbols/s without equalization.

The performance of the proposed lens antennas was also studied in a medium size room (24 m long by 11 m width, see [5]). For such dimensions, an aperture $\theta_M \sim 78^\circ$ was defined for both the BST and the MT lenses. The resulting power distribution is quite uniform for $h_b > 3$ m and the SDW can be as high as 80 ns, which means that a CBR of 90 Mbit/s can be achieved with the equalizer allowing also a significant tilting freedom of the MT. A similar figure can be achieved by using two pyramidal horns located in the two narrower walls of the room pointing towards each other and tilted downwards covering each one 2/3 of the room, see [1] [3]. This solution is more expensive to implement since it requires two BSTs, and eventually restricts more severely the MT freedom of movement.

Conclusions

The proposed shaped dielectric lens antennas were shown to produce sharp cells with uniform power distribution, and little restriction of MT movement. BST antenna height controls the cell radius and consequently the amount of side wall illumination. Experimental and simulation results indicate that these antennas outperform conventional solutions based on horns and biconics in terms of achievable CBR. Other lens combinations were tested leading to similar results.

References

- [1] J. Fernandes, "Transmission Capacity of a Broadband Wireless Radio Link", In these proceedings.
- [2] J. Fernandes, et. al., "Performance Evaluation of mm-wave Wide-Band Digital Radio Transmission", IEEE 2nd Symp. on Comm. and Vehicular Technology in the Benelux, Nov. 1994.
- [3] J. Fernandes, O. Afonso and J. Neves, "Impact of Antenna Set-up and Arrays on Mobile Radio Systems", ICUPC'95, Japan, Nov. 1995.
- [4] C.A. Fernandes, V. Brankovic, S. Zimmermann, M. Filipe and L. Anunciada, "Dielectric Lens Antennas for Wireless Broadband Communications", To appear in the special issue of the Wireless Personal Communications Journal on "Wireless Broadband Communications", 1998
- [5] J. Fernandes, P. Smulders and J. Neves, "Mm-wave Indoor Radio Channel Modelling vs. Measurements", Wireless Personal Communications Journal, Vol. 1, No 3, pp. 211-219, 1995.
- [6] J. C-I Chuang, "The Effects of Time Delay Spread on Portable Radio Communication Channels With Digital Modulation", IEEE J. On Selected Areas in Comm., Vol. SAC-5, June 1987, pp.879-889.

Statistical Behavior and Performance of Adaptive Antennas in Multipath Environments

Tien D. Pham, *Member, IEEE*

Mentor Graphics Corporation, 1001 Ridder Park Drive, San Jose, CA 95131-2314;

Tel: (408) 451-5529; Fax: (408) 451-5722; Email: tien_pham@mentorg.com.

Abstract

Adaptive antennas can behave very differently in multipath and non-multipath environments. In this paper, we study the behavior of adaptive antennas in multipath Rayleigh fading by investigating the statistical variation of the spatial signature of a signal incident upon an antenna array. A method for analyzing the multipath performance of adaptive antennas is presented. It can be shown that the antenna performance becomes less dependent on the angles of arrival (AOA) of the incident signals as their angular spreads and antenna element spacing increase. In general, antenna performance improves for increasing element spacing because fading reduction is more effective and interference suppression is possible even when the mean AOAs of desired and interfering signals coincide.

1. Introduction

Conventionally, adaptive antennas have been widely used in radar and satellite communications for interference suppression and signal tracking, where their performances are analyzed in a deterministic fashion because the AOAs of the incident signals are usually well-defined [1]. In wireless communications, adaptive antennas have been proposed for implementation at base stations to suppress co-channel interference and hence significantly increase the system capacity [2]. Because of local scattering around a mobile transmitter, the signal arrives at the base station via multipath propagation resulting in a variation of the branch signal phase and amplitude. Not only does this mean fading occurs at the antenna elements but also the phase relationship between the branch signals, which determines interference suppression, is randomized. Thus, the antennas performance must be analyzed as a statistical problem. There are roughly two main approaches for the implementation of adaptive antennas. The first one involves forming beams toward target transmitter and steering nulls toward interferers. This approach requires AOA information which can be estimated from the received signal, but it is not suitable for multipath environments for several reasons. First, there is no well-defined AOA for the signal and even if the multipath components of a signal can be resolved, their number can easily exceed the antennas degree of freedom so steering nulls toward all of them is impossible. Second, forming beams and steering nulls toward individual multipath components usually does not yield optimum signal-to-interference-plus-noise ratio (SINR) that can be achieved by the

array. The second approach is optimum combining which maximizes the SINR at the array output. By making use of a reference or other known characteristics of the target transmitter, its spatial signature as seen by the array can be identified and used to extract the desired signal while suppressing the interference as much as possible. Implicitly, an adaptive antenna that maximizes the SINR also reduces fading provided the branch signals are uncorrelated. In fact, in the absence of interference, an adaptive antenna maximizes the signal-to-noise ratio (SNR) which is exactly what a diversity antenna with maximal ratio combining does. Research on adaptive antennas has been focusing on developing algorithms with little attention paid to antenna design. It is worth noting that whatever algorithm may be invented, their SINR performance will be bound by that of optimum combining, which is determined by the array geometry, antenna element patterns and source distribution.

In this paper, we will study the multipath behavior of adaptive antennas with optimum combining, and present a method to analyze the statistical performance of adaptive antennas in Rayleigh fading.

2. Propagation model and spatial signature

A simple but useful model for multipath propagation [3] is shown in Fig. 1. The signal from a mobile transmitter arrives at a base station via direct line-of-sight path and reflections off the local scatterers surrounding the transmitters. The multipath waves are essentially confined in an sector with mean AOA ϕ_m and angular spread σ . Assuming the relative time

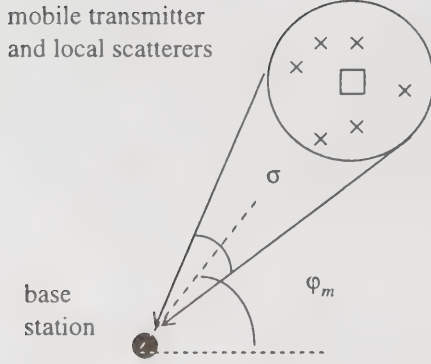


Fig. 1 Multipath propagation model.

delays for different paths are small compared to the inverse of the signal bandwidth, the received signal undergoes Rayleigh fading with random phase fluctuation.

Let an antenna array having M elements with the propagation vector

$$\mathbf{a}(\varphi_i) = [a_1(\varphi_i) \dots a_M(\varphi_i)]^T \quad (1)$$

for a plane arriving at angle φ_i in the horizontal plane, then the signal received at the antenna elements due to multipath propagation can be expressed in vector form as

$$\mathbf{y} = s(t)\mathbf{u} + \mathbf{n}(t) \quad (2)$$

where $s(t)$ is the transmitted signal,

$\mathbf{n}(t) = [n_1(t) \dots n_M(t)]^T$ is thermal noise vector

and $\mathbf{u} = [u_1 \dots u_M]^T$ is the spatial signature of the multipath signal as seen by the array, which is given by

$$\mathbf{u} = \sum_i \mathbf{a}(\varphi_i) r_i e^{-j\phi_i} \quad (3)$$

where r_i and ϕ_i are the amplitude and phase of the incident wave by the i th path.

For a large number of paths, the spatial signature is a complex Gaussian random vector with the covariance matrix

$$\mathbf{R} = \langle \mathbf{u}\mathbf{u}^H \rangle = \begin{bmatrix} R_{11} & \dots & R_{1M} \\ \vdots & \ddots & \vdots \\ R_{M1} & \dots & R_{MM} \end{bmatrix} \quad (4)$$

where R_{mn} is the covariance between the signals in the i th and j th branches. If the paths are independent, it can be shown that [4]

$$R_{mn} = \oint a_m(\varphi) a_n^*(\varphi) S(\varphi) d\varphi \quad (5)$$

where $S(\varphi)$ is the source distribution of the incident signal.

3. Variation of spatial signature

The SINR performance of an adaptive antenna depends on the spatial signature of the desired and interfering signals. For deterministic, single-path propagation, the spatial signature of a signal is proportional to the propagation vector of the incident wave as can be seen from eq. (3). For instance, the spatial signature of a signal impinging on a linear array of 3 uniform elements with spacing d is given by

$$\begin{aligned} \mathbf{u} &= r_1 e^{-j\phi} \mathbf{a}(\varphi) \\ &= c \begin{bmatrix} 1 & e^{-jkd \cos \varphi} & e^{-j2kd \cos \varphi} \end{bmatrix}^T \end{aligned} \quad (6)$$

where c is a proportionality constant, which can be ignored in the following analysis. In this case, the array performance clearly depends on the AOA of the incident signal. For $d = 0.5\lambda$ and a signal arriving at $\varphi = 45^\circ$, the components u_1 , u_2 and u_3 of the spatial signature are complex scalars with phases 0° , 127° and 255° , respectively.

For dynamic multipath propagation, the spatial signature varies randomly and it is unclear if the performance is AOA dependent. To find out, we could simulate the spatial signature variation by generating a large number of propagation paths with random amplitudes and phases, which are confined in an angular sector as shown in Fig. 1, then determining the spatial signature using eqs. (6) and (3). A better alternative is to calculate the covariance matrix using eq. (5) and then generate a set of correlated complex Gaussian random variables to represent the spatial signature. As an example, consider a multipath signal incident upon the above array with the mean AOA $\varphi_m = 45^\circ$ and angular spread σ .

Assume the incident waves are uniformly distributed in the angular sector, the covariance matrix is calculated for several values of d and σ . To illustrate the spatial signature variation, we generate 1,000 samples of the spatial signature in each case and plot its components u_1 , u_2 and u_3 on a set of complex planes as shown in Fig. 2. The average amplitudes of u_j are normalized to

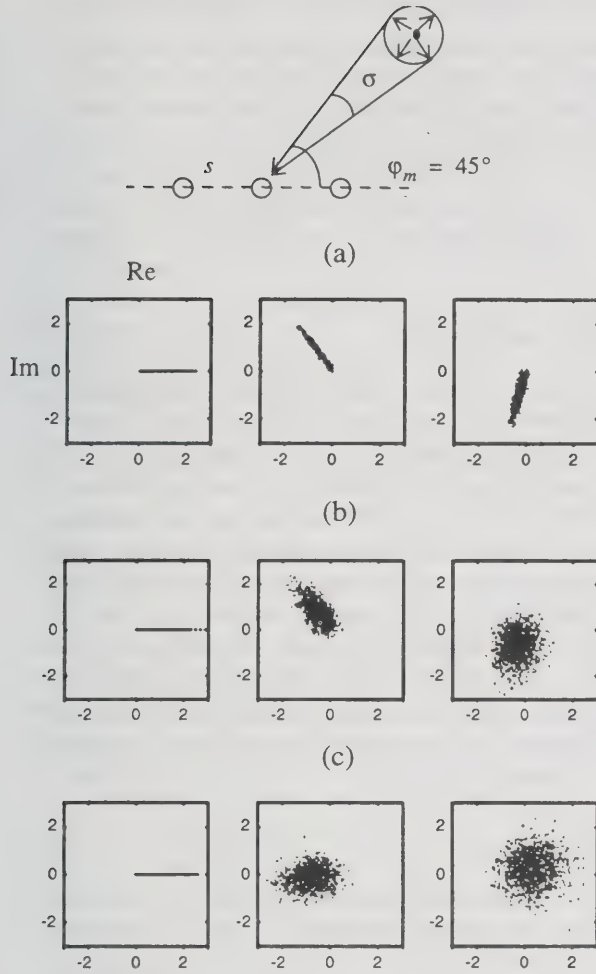


Fig. 2 Spatial signature cloud of a multipath signal with angular spread σ incident upon a three-element linear array with element spacing s at the mean angle $\varphi_m = 45^\circ$: (a) $\sigma = 5^\circ$, $s = \lambda/2$; (b) $\sigma = 30^\circ$, $s = \lambda/2$; (c) $\sigma = 5^\circ$, $s = 5\lambda$.

unity and their phases are shown relative to that of u_1 .

In general, we see that the amplitudes of u_j vary over a wide range and they take values close or equal zero, which means deep fades occur at all branches. For the relatively small spacing of $d = 0.5\lambda$ and angular spread $\sigma = 5^\circ$, the branch signals are highly correlated, especially between adjacent elements. Because of this, the relative phases between the elements do not change significantly. This is evident from the limited fluctuations of u_2 and u_3 phases from their mean values, which are the same as those in the deterministic case. Thus, the high correlations between the

branch signals have the following effects on the antenna performance. First, fades in the branches are correlated, which means the array cannot reduce fading very effectively. Second, the relative phases of the branch signals are essentially constant, which means interference suppression depends heavily on the AOAs of the incident signals. If the desired signal and interference have similar AOAs, the array cannot suppress the interference at any time. This is a well-known fact from conventional, deterministic analyses of adaptive antennas performance [1]. For increasing angular spread as in case (b), or increasing element spacing as in case (c), branch correlations are reduced and the relative phases undergo large fluctuations, which is evident from the significant extension of the spatial signature “clouds”. The low correlations mean fades occur more or less independently at different branches, thus the array can reduce fading more effectively. Furthermore, phase fluctuation means the array performance becomes less dependent on the AOAs of the incident signals, i.e. interference suppression is possible even if the desired signal and interference have the same mean AOAs. Even though phase fluctuation may occasionally degrade interference suppression, the output SINR rarely drops to extremely low values because simultaneous fades are unlikely.

4. Statistical performance

The multipath performance of an adaptive antenna can be determined using the result from the work of Pham and Balmain [5], who derived an approximate, closed-form expression for the cumulative distribution function (CDF) of the output SINR of an adaptive array in the presence of multiple interferers and correlated fades. Consider a linear array and a desired signal having the same propagation conditions as studied in cases (a)-(c) above. Furthermore, let an interfering signal with angular spread $\sigma_i = 5^\circ$ incident upon the array with the mean AOA $\varphi_{mi} = 90^\circ$. Assuming both the desired and interfering signals are 20dB above thermal noise, the CDF of the output SINR for the 3 cases are shown in Fig. 3. In case (a) we see that the output SINR is 5dB at 1% probability level meaning there is 1% chance that the output SINR is below 5dB. In cases (b) and (c), the output SINR is approximately 10dB at the same probability. Thus, the array performs better for large angular spread or increased element spacing.

To see how the array performance depends on the AOAs of the incident signals, let the mean AOA of the desired signal be fixed at 45° and that of the inter-

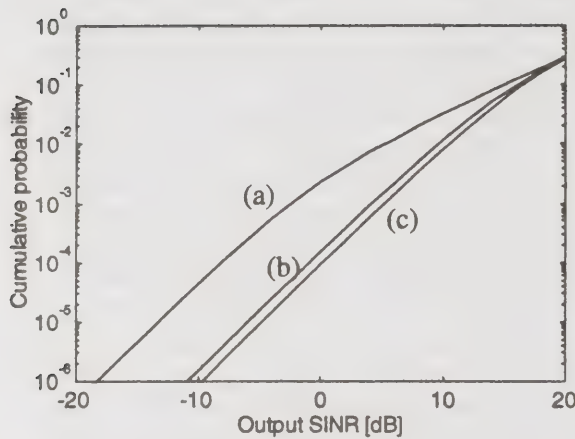


Fig. 3 The CDF of the output SINR for the linear array with same propagation conditions for the desired signal as shown in Fig. 2. The interference has a mean AOA $\phi_{mi} = 90^\circ$ and angular spread $\sigma_i = 5^\circ$ in all cases.

ference varying between -90° and 90° . The output SINR at 1% probability level for the above 3 cases is calculated and plotted in Fig. 4. In case (a), we see that

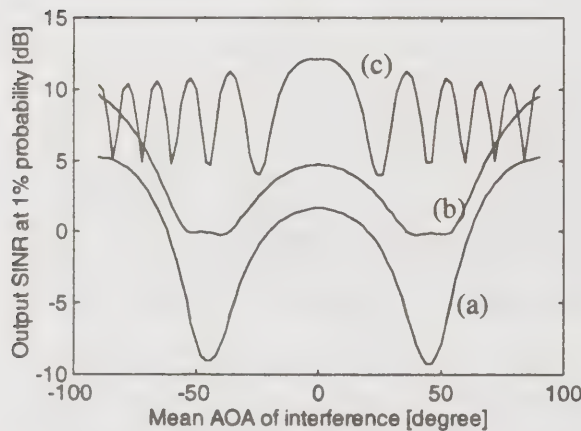


Fig. 4 1% output SINR performance of the linear array with same propagation conditions for the desired signal as shown in Fig. 2. The interference has an angular spread $\sigma_i = 5^\circ$ and mean AOA between -90° and 90° in all cases.

the SINR drops to -9dB for $\phi_{mi} = \pm 45^\circ$ where the spatial signatures of the desired and interfering signals have similar phases and hence interference suppression is not possible. In case (b), the SINR increases to 0dB for

$\phi_{mi} = \pm 45^\circ$ showing that the array performance is less dependent on the AOA of the interference. The performance is also better at other AOAs because fading reduction is more effective due to low branch correlations. In case (c), the SINR undergoes rapid but small variations. Overall the performance is better than in case (a) for all AOAs of the interference because both fading reduction and interference suppression are effective.

5. Conclusions

In this paper, we have studied the behavior of adaptive antennas and present a method to analyze their statistical performance in multipath environments. It was shown that the SINR performance becomes increasingly independent on the AOA of the received signal as the angular spread and element spacing increase. The results suggest that arrays with large element spacings should be used for base stations because both fading reduction and interference suppression are more effective than arrays having small element spacings. The presented method is general and can be used to analyze the performance of arrays with different geometries and element patterns.

6. References

- [1] R. T. Compton, Jr., *Adaptive Antennas: Concepts and Performance*, Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [2] J. Litva and T. K-Y. Lo, *Digital Beamforming in Wireless Communications*, Norwood, MA: Artech House, 1996.
- [3] W. C. Y. Lee, *Mobile Communication Engineering*, New York: McGraw-Hill, 1982.
- [4] R. G. Vaughan and J. Bach Andersen, "Antenna diversity in mobile communications," *IEEE Trans. Veh. Technol.*, vol. VT-36, no. 4, pp. 149-172, Nov. 1987.
- [5] T. D. Pham and K. G. Balmain, "Multipath performance of adaptive antennas in the presence of multiple interferers and correlated fadings," accepted for publication in *IEEE Trans. Veh. Technol.*

On Applications of Self-Phased Array Antennas to Mobile Communications

Sergey L. Loyka and Vladimir I. Mordachev
Belorussian State University of Informatics & Radioelectronics
P. Broyki Str. 6, Minsk 220027, Republic of Belarus
375-(0)17-2398994, e-mail: loyka@nemoc.belpak.minsk.by

Abstract

A simple method of substantial reduction (up to 1-2 orders) of power of base station and mobile unit transmitters used in mobile communication systems by the use of a self-phased array antenna which employs a pilot signal has been proposed. The increase in talk time and the reduction of battery weight by this method is also possible. Estimations of transmitted power reduction and examples of frequency assignment are given. Influence of electromagnetic environment on the self-phased array operation is discussed.

Key words: Self-phased array antenna, mobile communications

Introduction

Widespread utilisation of mobile communications systems which has lead to concentration of their transmitters over the limited territories (large populated and industrial areas) has dramatically aggravated electromagnetic and ecological compatibility problems. The radiated power of these systems' base stations can reach 5-10 kW (up to 320-350 W per channel), which can be ecologically dangerous in case their antennas are installed in residential districts as well as in the vicinity of industrial premises. Because of their comparatively large power, portable (0.5-5W) and mobile (5-50 W) units represent significant danger to their users. [1-3]. Approximately half the radiated power of a hand-held mobile phone is absorbed in a person's head [4]. According to certain data, the radiated power level of portable units which is acceptable from the medical point of view is 100-300 mW [5]. Apart from the ecological danger, a lot of base and mobile units in a mobile communications system over the limited area impairs the electromagnetic environment to a substantial extent, which causes many EMC problems. Besides, power consumption and, consequently, battery weight and capacity are very acute problems in mobile units design [6].

The Use of a Self-Phased Array Antenna

The above-mentioned problems can be fully solved by substantial reduction of the power radiated by mobile communications system transmitters. In this case the service areas of these systems may retain their size if a transition to the microcellular level is made as well as if sensitivity of the receivers is increased or

antennas gain is expanded. There can be no possible increase in sensitivity of receivers since this value is close to the limit (taking into consideration the state-of-art technological level and external noise level). It is hard to expand the gain of the mobile unit antenna because of the limitations on its weight and size. Transition to the microcellular level substantially increases cost and complexity of mobile communications systems and requires significant expansion of the radio frequency resource in order to ensure the intercellular communication.

The most promising way to retain the communication range of a system and size of its service and radio interference areas in this case is to use directional antennas for base stations so that the effective radiated power will be constant. Due to the random direction of a signal received by the antenna from a mobile unit it is necessary to employ the self-phased array antenna, which automatically forms the main lobe in the direction of the received signal [7-9]. This antenna type facilitates not only signal reception from any direction by its main lobe, but also signal reradiation in the same direction [7]. The theory and technology of these antennas were developed earlier owing to the fact that they are used for satellite communications systems where the exact positioning of an antenna to the satellite is required. There are several different types of self-phased array antennas. We will take into consideration the self-phased array antenna which employs the pilot signal, because of its simplest design. The typical block diagram of this antenna is represented at Fig.1. It includes array elements e_1-e_N , self-focusing modules u_1-u_N and a summer Σ . The self-focusing modules include filters tuned to the pilot and primary (information) signal frequency, a pilot signal

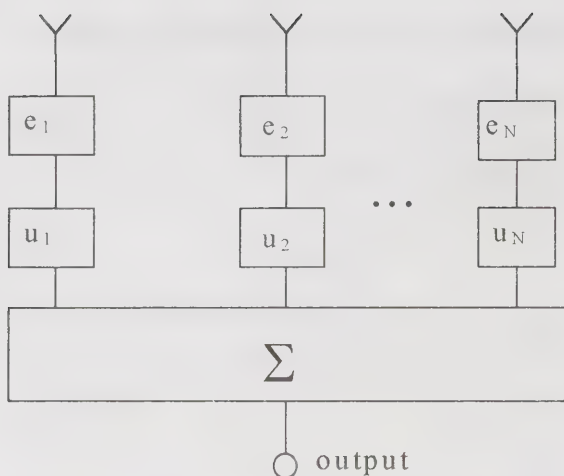


Figure 1. Block diagram of a self-phased array antenna. $e_1 - e_N$ - are array elements; $u_1 - u_N$ - are self-phasing units; Σ - is an adder.

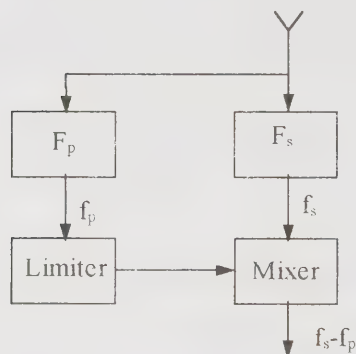


Figure 2. Block diagram of the self-focusing module. F_p, F_s - are filters tuned to the pilot signal frequency f_p and the primary signal frequency f_s .

limiter, and a mixer (Fig.2). The filter tuned to the intermediate frequency is connected after the summer. Phase shift compensation is carried out in the mixer during the multiplication of the pilot and primary signals. The signal phase at the intermediate frequency is identical at the output of all elements. In this case the main lobe direction always coincides with the direction of pilot signal reception. The operating principles of these antennas are described in detail in [7-9].

Let us estimate the advantage gained due to the reduction of power radiated by the base station transmitter, which can be achieved by using the self-phased array antenna. The constant character of a communication range (radius of a base station service

area) determines the constant character of the effective radiated power

$$P_e = G \cdot P_t = \text{const} \quad (1)$$

where P_t - is a transmitter power, G - is an antenna gain.

Hence it follows that in case the antenna is a non-directional one in the horizontal plane, for example, the dipole, the required transmitter power is

$$P_{t,d} = \frac{P_e}{G_d} = \frac{P_e}{1.64} \quad (2)$$

where $P_{t,d}$ - is the transmitter power for the dipole, G_d - is the gain of this dipole. The advantage due to the reduction of power radiated by the transmitter (admissible power reduction) is

$$K = \frac{P_{t,d}}{P_t} = \frac{G}{1.64} \quad (3)$$

where P_t - is the transmitter power for the self-phased array antenna; G - is the gain of this array. If the self-phased array is designed as a circular array antenna which consists of dipoles then, taking into account the shadowing effect, $G \approx 1.64 \cdot N$, where N - is the number of phase array elements (for the conversion factor in self-focusing modules equal to unity). Then the ratio (3) takes the form

$$K \approx N \quad (4)$$

Thus the use of the self-phased array antenna for the base station facilitates reduction of power radiated by its transmitter and a mobile unit transmitter by 10-20 dB (for $N=10 \div 100$). In this event the required transmitter power value for the cellular communications system will not exceed 0.01-0.1 W, and for the mobile unit paging system with active response - 0.05-0.1 W, which means that the most stringent sanitary requirements for the mobile communications system are satisfied without increasing the number of base stations. Additional advantages of using the self-phased array antenna are as follows:

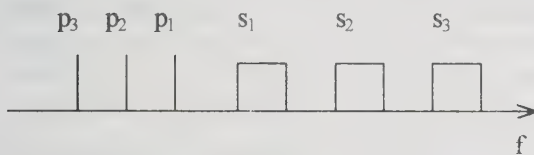
- The opportunity of increasing the continuous talk time of mobile units by 1-2 orders while preserving the capacity of batteries which encourages the further development of the radio paging with active response;
- The opportunity of decreasing the battery weight.
- The directional properties of the self-phased array antenna used for the base station make it possible to

considerably reduce the volume of the radio frequency resource actually used by the station (size of the radio-interference zone determined using probability-statistical criteria).

Further, we would like to note a number of possible limitations and peculiarities connected with using the self-phased array antenna for the mobile communications system:

- In the case of multichannel system structure it is necessary to use a pilot signal which has a different frequency for each channel. In this event the pilot signals of separate channels can be distributed either within a separate bandwidth, or within the guard band between the neighbouring channels (Fig.3).

a) p - pilot signal; s - primary signal.



b)

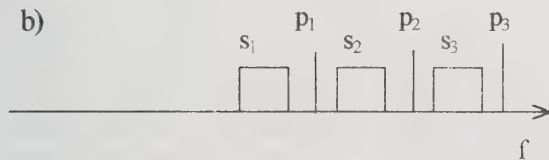


Figure 3. Frequency distribution of primary and pilot signals.

- The pilot signal is a harmonic one, and it is possible to carry out its filtering in the self-focusing module by the narrowband filter whose pass band can be by 2-4 orders less than the pass band of the filter for the primary signal and can be chosen taking into account the danger of dynamic effects in the system. Hence, on tentative estimations, the pilot signal power can be by 20-40 dB less than the primary signal power, which is very important for ensuring intrasystem EMC of base stations and mobile units which comprise the mobile communications system.
- The use of the self-phased array antenna will raise the cost of base stations and mobile units as well as the corresponding operational costs. However, on tentative estimations, this raise will be rather small.
- The use of adaptive phase array antennas for mobile communications systems [17-19] is complicated

owing to the sharp deterioration of their characteristics when the number of mobile units which work simultaneously is close to the number of adaptive phased array elements, complexity of the adaptation procedure and substantial amount of adaptation time, and complexity of an adaptive array design.

Influence of Electromagnetic Environment on Self-Phased Array Operation

Electromagnetic environment can essentially influence the self-phased array operation due to the active elements which are present in its structure. This influence can manifest itself in two main ways: (1) distortion of the array pattern due to multipath propagation (reflections) or jamming signals, and (2) generation of the intermodulation products at the array output ("intermodulation lobes").

Presence of reflections (multipath propagation) or jamming (interference) signals at the pilot signal frequency can lead to significant deterioration of a self-phased array operation. Both the pilot signal channel and the primary (information) signal channel can substantially influence on the self-phased array operation and must be taken into account during the analysis [10-12]. In the case of the threshold amplitude characteristic of the pilot-signal channel, directivity of self-phased array under conditions of multipath propagation can be written as follows [10-12]:

$$G = G_p \cdot G_s \cdot G_0 \quad (5)$$

where G_0 - is the array directivity in the absence of reflections; G_p - is reduction in directivity due to cutting off some channels (influence of reflections by way of the pilot signal channel); G_s - is reduction in directivity owing to variations in amplitude and phase of the primary signal (influence of reflections by way of the primary signal channel). As the detailed analysis shows, the average reduction in directivity under normal operation conditions is no more than 6 dB, but more larger reduction is possible in some specific cases.

Presence of intermodulation lobes in a self-phased array can substantially decrease its interference immunity (or, equivalently, increase its susceptibility to interference) and distort output signal. It should be noted that the array pattern at an intermodulation frequency may be quite different from that at the working frequency. A detailed analysis of intermodulation lobes in active array antennas can be found in [13-16].

Conclusion

A self-phased array antenna is very promising for the use in mobile communications. Its design is very simple: it performs self-phasing without phase shifters and adaptive control electronics. The large progress in the field of microwave IC technology makes its production very simple and cheap.

By and large, the use of the self-phased array antennas in mobile communications systems requires for the development of a special standard which would regulate the use of radio frequency resource and the ensuring of EMC for base stations and mobile units.

References

- [1] I. Shtrikman, and H. Fruchting, 'Numerical Determination of the Performance and of the Electromagnetic Compatibility of Antennas for Hand Held Mobile Telephones', *Proceedings of 13th Wroclaw Int. Symp. on EMC*, pp.189-193, Wroclaw, Poland, 1996.
- [2] M. Fischetti, 'The cellular phone scare', *IEEE Spectrum*, pp.43-47, June 1993.
- [3] H.J. Mikolajczyk, and M. Kamedula, 'Direct Experimental Evidence of Microwave Influence on the Central Organs of Endocrine System', *Proceedings of 13th Wroclaw Int. Symp. on EMC*, pp.130-132, Wroclaw, Poland, 1996.
- [4] Y. Rahmat-Samii and K.W. Kim, Antennas and Human in Personal Communications: Applications of Modern EM Computation Techniques, 12th Int. Confer. on Microwaves and Radar (MIKON'98), Krakow, Poland, May 20-22, 1998. pp.36-55.
- [5] Sanitary norms for admissible levels of the physical factors concerning the use of the consumer goods in household conditions. - *Sanitary regulations 9-29-95* (Russian Federation. 2.1.8.042 -96).
- [6] D.C. Cox, 'Wireless Personal Communications: A perspective'. In *"The Mobile Communications Handbook"*, Editor-in-Chief J.V. Gibson, CRC Press, 1996.
- [7] P.V. Brennan, 'An Experimental and Theoretical Study of Self-Phased Arrays in Mobile Satellite Communications', *IEEE Trans. on Antennas and Propag.*, Vol.37, No.11, pp. 1370-1376, 1989.
- [8] M.J. Withers et al., 'Self-focusing receiving array', *Proc. IEE*, Vol. 112. No.9, pp.1683-1688, Sep 1965
- [9] 'Antenna arrays: methods of calculation and designing. A survey of foreign sources', Ed. by L.S. Benenson, Moscow, Sovetskoe radio, 1966.
- [10] S.L. Loyka, Influence of Pilot Signal on Directivity of Self-Phased Arrays Under Conditions of Multipath Propagation, *IEEE Trans. on EMC*, vol.40, No.1, pp.12-18, 1998.
- [11] S.L. Loyka, On Operation of Self-Phased Arrays Under Conditions of Multipath Propagation, *Int. Conf. on Microwave & Radar (MIKON-98)*, Poland, Krakow, 1998, pp.535-539.
- [12] Loyka S.L. 'Directivity of Self-Phased Arrays in Conditions of Multipath Propagation,' *Proceedings of the Seventh International Conference on HF Radio Systems and Techniques*, East Midlands Conference Centre, Nottingham, UK, 7-10 July, 1997.
- [13] S.L. Loyka, Characteristics of receiving intermodulation channel of active array antennas, *International Journal of Electronics*, Vol.80, No. 4, pp. 595-602, 1996.
- [14] S.L. Loyka, Interference Immunity of Active Array Antennas in Rigid Electromagnetic Environment, *Proceeding of the XI International Microwave Conference "MIKON-96"*, Poland, Warsaw, 1996, vol.1, pp.57-61.
- [15] S.L. Loyka, Intermodulation receiving channels in active array antennas, *Izvestia Vuzov. Radioelectronica* (Radioelectronics and Communications Systems, Allerton Press Inc.), vol.39, N 2, 1996. pp.68-74.
- [16] S.L. Loyka, Conditions of existence of main lobes in two-signal spatial selectivity characteristic of active array antennas, *Izvestia Vuzov. Radioelectronica* (Radioelectronics and Communications Systems, Allerton Press Inc.), vol.39, N 6, 1996, pp.3-9.
- [17] R.A. Monzingo and T.W. Miller, *'Introduction to Adaptive Arrays'*. John Wiley & Sons, New York, 1980.
- [18] B. Widrow and S.D. Stearns, *'Adaptive Signal Processing'*. Prentice-Hall, Inc., Englewood Cliffs, 1985.
- [19] D. Nowicki, J. Ronmeliotis, 'Smart Antenna Strategies', *Mobile Communication International*, April 1995, pp. 53-56.

A New Real Time Radio Frequency Direction Finding Algorithm for Gaussian and non-Gaussian Noise Environments

W. Featherstone and H. J. Strangeways

Institute of Integrated Information Systems,
School of Electronic and Electrical Engineering,
University of Leeds, Leeds LS2 9JT, UK

tel: +44 (0) 113 233 2075

fax: +44 (0) 113 233 2032

email: eenwf@electeng.leeds.ac.uk

Abstract

In this paper, a new superresolution direction finding (SRDF) algorithm for multiple incident radio waves is proposed. Superresolution methods enable resolution of signals separated by less than the natural beamwidth of the array. This ability enables the algorithms to separate the closely spaced signals encountered in a multipath environment. The algorithm is termed Loaded Capon and is shown to be capable of operating on data sets containing a limited number of data points. The new algorithm is shown to be robust in both Gaussian and non-Gaussian noise environments. Simulated and measured data, recorded on a multi-channel direction finding system, are used to demonstrate the algorithm's superior performance robustness over both the standard MVE and eigen-based techniques such as MUSIC.

Key Words: Superresolution Direction Finding, Array Signal Processing

1. Introduction

In the derivation and analysis of most radio frequency superresolution direction finding algorithms, an idealised noise model is assumed. The assumptions are made that the noise process observed at each element of a sensor array is Gaussian having zero mean and variance σ^2 , and, in addition, each process is temporally and spatially uncorrelated. In real radio communication channels, it is often found that the dominant source of interference contains a significant noise component that is impulsive, indicating the probability of large interference levels. In addition, the noise may not be independent and identically distributed (iid) due to correlation of the noise across the array. For example, correlation may arise due to propagational effects or may be induced by narrowband filtering of uncorrelated noise. In such situations, the impulsive events may be spread over many data samples. For these real situations there is a severe degradation in performance for most SRDF techniques.

In this paper we propose a new SRDF algorithm, termed Loaded Capon, that is more robust to non-Gaussian environments than the MVE [1] algorithm upon which it is based and also eigenbased algorithms like MUSIC [2]. It is shown that, as the

number of data samples is reduced, the performance of the standard MVE degrades. The degradation is shown to be caused by the spread of the noise subspace eigenvalues of the covariance matrix. The new method overcomes the aforementioned problems by loading the leading diagonal of the covariance matrix with a fraction of the total power contained within it. Carlson [3] proposed a similar form of covariance loading for adaptive beamforming when using sampled matrix inversion. Results produced through simulation and measurement are used to quantify the performance gains through use of the new method. The performance of Loaded Capon is also compared with the eigenbased method MUSIC, where it is found to have comparable performance in the idealized case, but far superior performance in correlated non-Gaussian environments. The benefits are obtained without the increase in computational complexity incurred by employing methods such as [4][5] that attempt to whiten the noise received at the array or methods such as MUSIC that are based on an eigendecomposition.

This paper addresses the problem of limited data sets, that is, when only a limited number snapshots are available for the formation of the covariance matrix, such as, in the case of real time channel

evaluation (RTCE). Further, by limiting the sampling time, it is possible to determine DOAs in time varying channels in a shorter time interval. This is particularly advantages when frequency hopping is being utilized. The advantages of Loaded Capon are most marked for limited data sets as it is found that, as the number of snapshots increases in a stationary channel, the performance of the ordinary MVE approaches that of Loaded Capon.

The paper is organized as follows. First in Section 2 the theory behind Loaded Capon is introduced. This includes the theoretical treatment of why the Loaded Capon algorithm outperforms the original MVE. In Section 3 the results obtained through simulation and measurement are given and finally in Section 4 conclusions are drawn from the results presented.

2. Theory

The MVE method determines the Direction of Arrival (DOA) by evaluating the following spectral function

$$P_c(\theta) = \frac{1}{\mathbf{a}(\theta)^H \mathbf{R}^{-1} \mathbf{a}(\theta)} \quad (1)$$

The subscript H represents the hermitian transpose operator and $\mathbf{a}(\theta)$ is the steering vector describing the transfer function between DOA θ and the output of the array. \mathbf{R} , the estimated covariance matrix, is defined as

$$\mathbf{R} = \frac{1}{K} \sum_{i=1}^K \mathbf{x}_i \mathbf{x}_i^H \quad (2)$$

where K is the number of samples and \mathbf{x}_i is the i th complex data vector recorded by the receiving array.

The covariance matrix can be rewritten as

$$\mathbf{R} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H \quad (3)$$

where

$$\mathbf{\Lambda} = \text{diag}(\lambda_1 \lambda_2 \dots \lambda_N)$$

$$\mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_N]$$

$\lambda_1 \lambda_2 \dots \lambda_N$ are the eigenvalues of \mathbf{R} and $\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_N$ are the corresponding eigenvectors.

By use of the Matrix Inversion Lemma, it can be shown that to within a scaling factor of $1/\lambda_{\min}$

$$\mathbf{R}^{-1} = \mathbf{I} - \sum_{i=1}^N \mathbf{v}_i \frac{\lambda_i - \lambda_{\min}}{\lambda_i} \mathbf{v}_i^H \quad (4)$$

where λ_{\min} is the smallest eigenvalue.

The MVE algorithm can now be evaluated as,

$$\frac{1}{\mathbf{a}(\theta)^H \mathbf{R}^{-1} \mathbf{a}(\theta)} = \frac{1}{\mathbf{a}(\theta)^H \mathbf{a}(\theta) - \sum_{i=1}^N \frac{\lambda_i - \lambda_{\min}}{\lambda_i} |\mathbf{a}(\theta)^H \mathbf{v}_i|^2} \quad (5)$$

Inspection of the right hand side of the above equation shows that each eigenvector contributes to the output in the proportion to $(\lambda_i - \lambda_{\min})/\lambda_i$. In the case of infinite samples every eigenvalue in the noise subspace has an eigenvalue of λ_{\min} . Thus since $(\lambda_i - \lambda_{\min})/\lambda_i = 0$, the noise subspace makes no contribution to the output in the infinite sample case.

In the finite sample case, the noise subspace eigenvalues are spread, i.e. $\lambda_{m+1} > \lambda_{m+2} > \dots > \lambda_{\min}$, where m is the rank of the signal subspace. Further, if the noise in the data samples is not statistically independent, the noise eigenvalues may also be of reduced magnitude. In this case $(\lambda_{m+i} - \lambda_{\min})/\lambda_{m+i} \neq 0$ (providing $\lambda_{m+i} \neq \lambda_{\min}$); hence, the noise subspace is corrupting the output.

If the sample support for the noise subspace can be made to look infinite, the MVE algorithm should give similar performance to MUSIC. The new method achieves this by loading the leading diagonal of the covariance matrix with a fraction of the total power contained within the covariance matrix. The consequence of diagonally loading the covariance matrix is to add a constant to each eigenvalue, which has the effect of reducing the $(\lambda_{m+i} - \lambda_{\min})/\lambda_{m+i}$ term, thus reducing how much the noise subspace corrupts the output. Providing the signal subspace eigenvalues are much larger than the loading level, the effect on the signal subspace's contribution to the output is minimal.

In addition the spread of the eigenvalues may give rise to an ill conditioned covariance matrix. In this case, the diagonal loading will help in the inversion of the covariance matrix.

The resulting algorithm is more robust to inherent modeling errors of the covariance matrix. The modified covariance matrix, \mathbf{R}_L , used in the new method is defined as

$$\mathbf{R}_L = \mathbf{R} + \alpha \mathbf{I} \quad (6)$$

where \mathbf{I} is the identity matrix and α is the loading factor. Simulation carried out by the authors suggest that an appropriate amount of loading is $\alpha = \text{Tr}[\mathbf{R}]/100N$. $\text{Tr}[\cdot]$ is the trace operator, which

extracts and sums the elements of the leading diagonal of the matrix.

3. Results.

To assess the performance of the new algorithm, simulations were carried out using a 10 element uniform linear array with a half-wavelength inter-element spacing. In all simulations, two uncorrelated signals are incident on the array at 84.6° and 95.4° , respectively. The variance of only one of the estimated direction of arrivals (DOAs) is shown for clarity. The number of snapshots used in the formation of the covariance matrix is altered to assess the resulting effect on Loaded Capon, MVE and MUSIC. White Gaussian noise is present at each antenna, with a signal-to-noise ratio (SNR) of 12 dB. It is seen from Figure 1 that the variances of the estimated DOAs made by Loaded Capon are considerably lower than the standard MVE. The benefits of using Loaded Capon increase as the number of samples available for the formation of the covariance matrix decreases. It will be seen that MUSIC only outperforms MVE and Loaded Capon in this white noise environment.

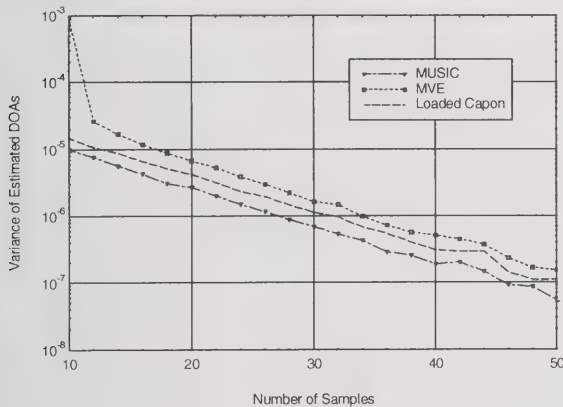


Figure 1. Comparative performance of the algorithms as the number of samples is increased.

The performance of the algorithms is then evaluated for a non-ideal noise environment. To introduce correlation into the noise process we adopt the approach of Sadler [6] that allows correlation in both the nominal background and impulsive components. In Figure 2, the performance of the algorithms is shown when both the Gaussian and impulsive components are passed through an auto-regressive (AR) filter of order one. An AR coefficient of -0.9 was chosen for both processes. The signal-to-impulse ratio (SIR) is set at -20 dB. Loaded Capon shows better performance than MVE regardless of the SNR. The performance of MUSIC is severely degraded by introducing non-ideal noise.

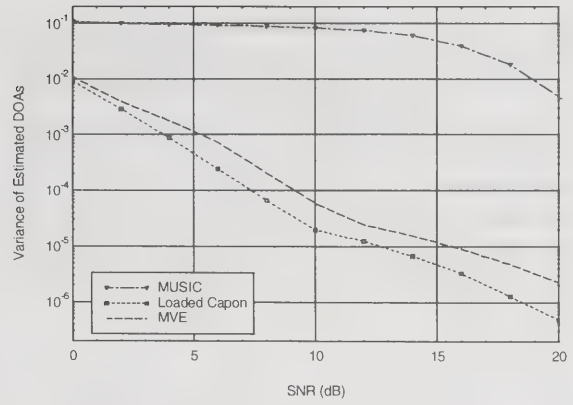


Figure 2. Comparative performance of the algorithms in correlated Gaussian plus Impulsive Noise.

Finally, Figure 3 shows the performance of Loaded Capon for different simulation scenarios using 20 samples in the formation of the covariance matrix. It is only for low SNRs that the variance of the DOA estimates deviates significantly from the purely Gaussian case. Consequently, it is seen that Loaded Capon is robust to deviations from the case of iid Gaussian noise.

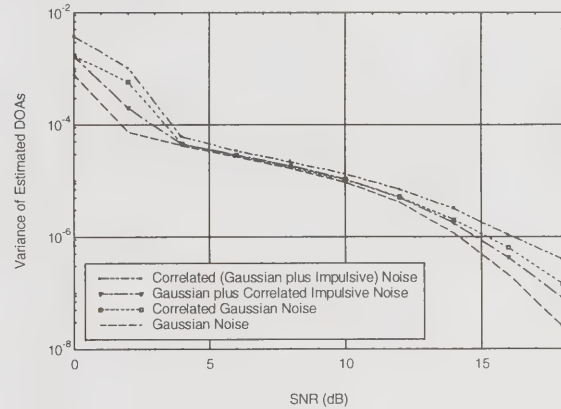


Figure 3. Performance of Loaded Capon in every scenario considered.

In addition to the simulated data measured data was recorded using a multi-channel direction finding system. To demonstrate the performance of the new algorithm the accuracy of its AOA estimates was compared against those made by the original MVE and MUSIC. The algorithms were used to determine the AOA of data received on a path from Hilversum (Radio Netherlands), to a receiver located near Hamburg, Germany. The true angle of arrival is known to be 244 degrees, the transmission frequency was 5.955MHz and the data was recorded at 10 GMT on 8 September 1994. The data sets consist of sixteen samples taken every $12\mu\text{s}$, with is

a one second spacing between sets. Figure 4 presents the comparison between the original MVE and the new Loaded Capon algorithm and shows that, for the majority of the sets, the MVE fails to correctly resolve AOA, whereas the Loaded Capon fluctuates closely around it. Figure 5 gives the results for the same data as for Figure 4, but here the comparison is for MUSIC against the Loaded Capon. The new algorithm slightly outperforms MUSIC on this data set.

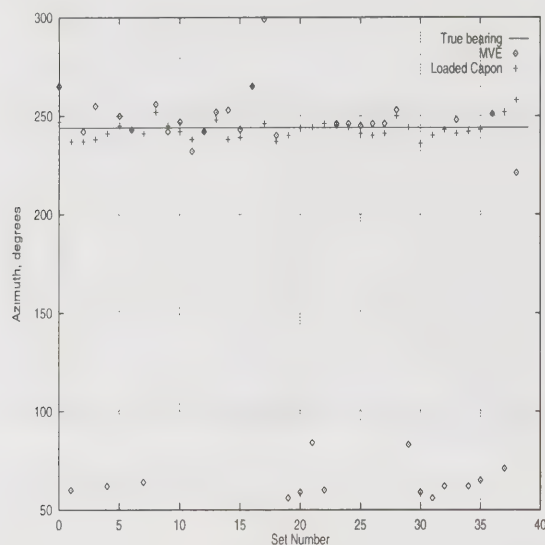


Figure 4. Plot showing the comparison of the estimated azimuth determined by the MVE and the Loaded Capon algorithms.

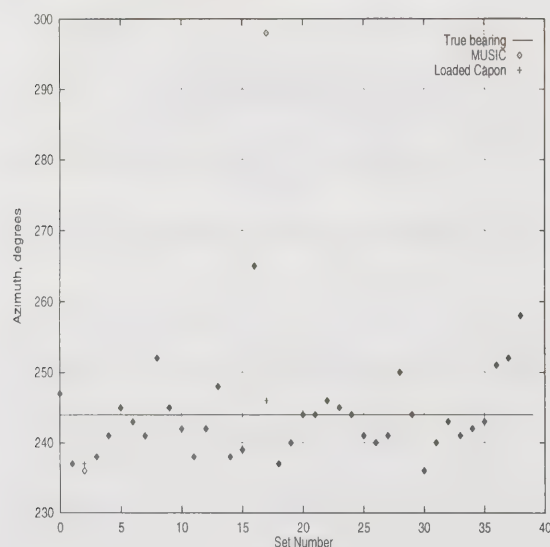


Figure 5. Plot showing the comparison of the estimated azimuth determined by the MUSIC and the Loaded Capon algorithms.

4. Conclusions

The problem of inadequately estimating the covariance matrix in the limited sample case has been highlighted when employing the MVE algorithm. In addition the detrimental effects on the MVE algorithm in coloured non-Gaussian environments have been shown, in particular when there is an impulsive component present in the observed data snapshot. It has been shown that the Loaded Capon algorithm, in which the leading diagonal of the covariance matrix is loaded can overcome these difficulties. The results presented show the improvement in performance (with no increase in complexity) of the Loaded Capon algorithm compared with the standard MVE for both the simulated scenarios and the measured data. Further, in comparison with the MUSIC algorithm it is shown the new method has reduced complexity with comparable bearing accuracy in the purely Gaussian case, but far superior performance to MUSIC in the coloured non-Gaussian environment.

5. References

- [1] Capon J, "High-Resolution Frequency-Wavenumber Spectrum Analysis" *IEEE Proc.*, vol. 57, pp 1408-1419, 1969.
- [2] Schmidt R O, "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans AP*, vol. 34, pp 276-280, 1986.
- [3] Carlson B D, "Covariance Estimation and Diagonal Loading in Adaptive Arrays", *IEEE Trans. AES*, vol. 24, pp 397-401, 1988.
- [4] Friedlander B and Weiss A J, "Direction Finding Using Noise Covariance Modeling", *IEEE Trans. SP*, vol. 43, pp 1557-1567, 1995.
- [5] Le Cadre J P "Parametric Methods for Spatial Signal Processing in the Presence of Unknown Colored Noise Fields", *IEEE Trans ASSP*, vol. 37, pp 965-983, 1989.
- [6] Sadler B M, "Detection in Correlated Impulsive Noise Using Fourth-Order Cumulates", *IEEE Trans. SP*, vol. 44, pp 2793-2800, 1996.

Harmonic Retrieval in Colored Non-Gaussian Noise Using Cumulant and Autocorrelation

Yan Zhang and Shu-Xun Wang

Department of Electrical Engineering, Jilin University of Technology

Changchun 130025, P.R.China

E-mail : wsx@jut.edu.cn

Abstract

This paper proposes a higher order cumulant-based approach to harmonic retrieval in non-Gaussian ARMA noise. Hilbert transform is used to transform the real measurements into their complex counterpart, then a kind of elaborately defined fourth-order cumulant is employed to identify the AR parameters of the non-Gaussian noise. After prefiltering the noisy measurements with the identified AR polynomial, SVD-TLS method can be applied to retrieve harmonics. This method releases the asymmetrically-distributed assumption of the non-Gaussian noise and is also efficient when there is quadratic phase coupling in harmonics. Simulation examples are presented to support the result of this paper.

Key words: harmonic retrieval, non-Gaussian colored noise, fourth-order cumulant, Hilbert transform

Introduction

The harmonic retrieval (HR) problem involves estimating the parameters of a harmonic signal observed in noise. It is common in signal processing and occurs in fields such as radar, sonar, and telemetry. Considerable research efforts have been witnessed during recent years. Most HR approaches either assumed white noise and utilized second-order statistics, or assumed colored Gaussian noise and employed the cumulant-based methods. Since colored non-Gaussian noise environments are frequently related to sonar systems and signal detection[1], more recently harmonic retrieval in colored non-Gaussian noise with asymmetrical distribution was studied in [4,5] with the assumption that no quadratic phase coupling occurs.

But quadratic phase coupled harmonics may be encountered in practice such as in oceanography, seismology and EEG analysis, while symmetrically distributed non-Gaussian noise arises in a variety of applications, see [6] and reference therein for detail. The purpose of this paper is to retrieve harmonics which may be quadratic phase coupled in colored linear non-Gaussian noise with no restriction on its distribution. We show that the complex counterpart of the real noisy observations can be obtained through Hilbert transform. Some particularly cumulant definition is proved suitable for estimating noise characteristics. Then the new prefiltering method is proposed. The key idea is that the cumulants of the non-Gaussian noise can be estimated in the presence of har-

monics, consequently the AR polynomial of the non-Gaussian noise is recovered, then these estimated noise characteristics are used to partially whiten (make MA) the noise. Finally conventional method can be applied to retrieve the harmonics. Experimental results illustrate the performance of the new approach.

Harmonics in Non-Gaussian Noise: Signal Model

The discrete-time signal-plus-noise model is given by

$$y(n) = s(n) + w(n) \quad (1)$$

where $s(n)$ is a real-valued harmonic signal given by

$$s(n) = \sum_{i=1}^p a_i \cos(\omega_i n + \varphi_i) \quad (2)$$

in (2) a_i 's and ω_i 's are constants while φ_i 's are uniformly distributed over $[0, 2\pi)$. The problem of interest is to estimate p and ω_i 's using just the noisy observations $y(n)$, $n = 1, \dots, N$.

ASSUMPTION The noise $w(n)$ is modeled as an ARMA(n_b, n_d) process given by

$$B(q)w(n) = D(q)e(n) \quad (3)$$

where $B(q) = \sum_{j=0}^{n_b} b(j)q^{-j}$, $D(q) = \sum_{j=0}^{n_d} d(j)q^{-j}$ and

$q^{-j}e(n) = e(n-j)$. The ARMA process is assumed exponentially stable and free of pole-zero cancellations, is causal but may be nonminimum phase.

REMARK Contrasting with the reported studies, there is no assumption that no quadratic phase coupling occurs, and the additive non-Gaussian noise $w(n)$ may be either asymmetrically or symmetrically distributed.

Estimation Of The Non-Gaussian Noise Characteristics

Quadratic phase coupling may incur in (2). Namely, there may exist integers $l_1, l_2, l_3 \in [1, p]$ such that both $\omega_{l_1} + \omega_{l_2} = \omega_{l_3}$ and $\varphi_{l_1} + \varphi_{l_2} = \varphi_{l_3}$ are satisfied. When the harmonics are quadratic phase coupled and/or the noise is symmetrically distributed, the noise characteristics can't be determined from both the third-order and fourth-order cumulants of the real noisy signal. Hence we transform the real observations into their complex counterpart. It is well known that for complex process there are many different cumulant definitions depending on the position of conjugations. Among them some particular definition, which is zero for coupled or uncoupled harmonics while is nonzero for symmetrically or asymmetrically distributed noise, may be chosen to estimate the noise characteristics.

Given a real process $x(t)$, its complex counterpart is

$$\dot{x}(t) = x(t) + j\bar{x}(t) \quad (4)$$

where $\bar{x}(t)$ is the Hilbert transform of $x(t)$. $\dot{x}(t)$ can be seen as the output of a linear model with the input $x(t)$, the impulse response of this linear model is

$$h_1(t) = \delta(t) - \frac{1}{j\pi t} \quad (5)$$

Obtain the complex counterpart of the real noisy signal $y(n)$ in (1) using Hilbert transform

$$\tilde{y}(n) = y(n) + j\bar{y}(n) = \tilde{s}(n) + \tilde{w}(n) \quad (6)$$

where $\tilde{s}(n)$ is the complex-valued harmonic signal

$$\tilde{s}(n) = \sum_{i=1}^p a_i \exp(j(\omega_i n + \varphi_i)) \quad (7)$$

$\tilde{w}(n)$ is the complex non-Gaussian noise. $\tilde{y}(n)$ can be depicted by the block diagram of Fig.1.

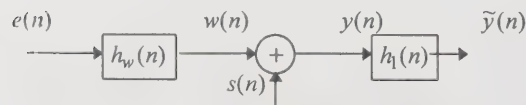


Fig.1. The block diagram which depicts the complex observations $\tilde{y}(n)$.

We show that the noise AR order n_b and AR coefficients $b_i, i = 1, \dots, n_b$ can be determined in the presence of the quadratic phase coupled or uncoupled harmonics with no restriction on the non-Gaussian noise distribution.

Definition 1: The fourth-order cumulant of a complex process $\tilde{x}(n), c_{4\tilde{x}}(\tau_1, \tau_2, \tau_3)$ is defined as

$$c_{4\tilde{x}} = \text{cum}(\tilde{x}(n), \text{Re}(\tilde{x}(n + \tau_1)), \tilde{x}(n + \tau_2), \tilde{x}(n + \tau_3))$$

where $\text{Re}(\tilde{x}(n + \tau_1)) = x(n + \tau_1)$ is the real part of $\tilde{x}(n + \tau_1)$. It is obvious that

$$c_{4\tilde{x}} = \frac{1}{2} (\text{cum}(\tilde{x}(n), \tilde{x}^*(n + \tau_1), \tilde{x}(n + \tau_2), \tilde{x}(n + \tau_3)) + \text{cum}(\tilde{x}(n), \tilde{x}(n + \tau_1), \tilde{x}^*(n + \tau_2), \tilde{x}(n + \tau_3))) \quad (8)$$

Then we can easily obtain the result that according to definition 1 $c_{4\tilde{x}}$ is zero no matter whether the harmonics are quadratic phase coupled, whereas $c_{4\tilde{w}}$ is nonzero as long as $w(n)$ is non-Gaussian.

Theorem 1: The fourth-order cumulants of $\tilde{w}(n)$ in (6) according to definition 1 satisfy the following higher order Yule-Walker equation

$$\sum_{i=0}^{n_b} b(i) C_{4,\tilde{w}}(\tau_1 - i, \tau_2, \tau_3) = 0 \quad \tau_1 > n_d \quad (9)$$

Proof: The impulse response of the ARMA model in the Assumption is $h_w(n)$, then

$$\text{Re}(\tilde{w}(n)) = w(n) = \sum_{\tau} h_w(\tau) e(n - \tau)$$

let $h(n) = h_w(n) * h_1(n)$ and from definition 1, we have

$$\begin{aligned} C_{4,\tilde{w}}(\tau_1, \tau_2, \tau_3) &= \text{Cum}(\sum_{i_0} e(i_0) h(l - i_0), \sum_{i_1} e(i_1) h_w(l + \tau_1 - i_1), \\ &\sum_{i_2} e(i_2) h(l + \tau_2 - i_2), \sum_{i_3} e(i_3) h(l + \tau_3 - i_3)) \\ &= r_{4,e} \sum_{n=0}^{\infty} h(n) h_w(n + \tau_1) h(n + \tau_2) h(n + \tau_3) \end{aligned}$$

$$\text{then } \sum_{i=0}^{n_b} b(i) C_{4,\tilde{w}}(\tau_1 - i, \tau_2, \tau_3)$$

$$= r_{4,e} \sum_{n=0}^{\infty} h(n) (\sum_{i=0}^{n_b} b(i) h_w(n + \tau_1 - i)) h(n + \tau_2) h(n + \tau_3)$$

$$\text{notice } \sum_{i=0}^{n_b} b(i) h_w(n + \tau_1 - i) = d(n + \tau_1), \text{ where } d(i) = 0$$

if $i > n_d$ or $i < 0$. Consequently when $\tau_1 > n_d$, the left-hand side of (9) is identical with zero. \square

From Theorem 1 the AR order and AR parameters can be determined by Giannakis and Mendel[3].

It is well known that the fourth-order cumulant of a complex process has 2^4 different definitions depending on the position of conjugations. The result can be obtained that none of these cumulant definitions alone satisfies the Yule-Walker equation in (9), thus they all can't be used to estimate the noise AR order and parameters. See detail in Appendix.

Prefiltering Method For Harmonic Retrieval

Having determined the noise AR parameters $b(i)$ $i = 1, \dots, n_b$, multiplying both sides of (1) by $B(q)$ and then using (3), we have

$$\begin{aligned} B(q)y(n) &= B(q)s(n) + B(q)w(n) \\ &= B(q)s(n) + D(q)e(n) \end{aligned} \quad (10)$$

$$\text{Denote } \tilde{y}(n) = B(q)y(n) = \sum_{i=0}^{n_b} b(i)y(n-i)$$

$$\tilde{s}(n) = B(q)s(n) = \sum_{i=0}^{n_b} b(i)s(n-i)$$

$$v(n) = D(q)e(n) = \sum_{i=0}^{n_d} d(i)e(n-i)$$

then (10) can be rewritten as

$$\tilde{y}(n) = \tilde{s}(n) + v(n) \quad (11)$$

We refer to $\tilde{y}(n)$ as the filtered output process and $\tilde{s}(n)$ as the filtered harmonic signal. It is noted that $v(n)$ is a purely $MA(n_d)$ process, we have

$$R_{\tilde{y}}(m) \equiv R_{\tilde{s}}(m), \quad m > n_d \quad (12)$$

It is well known that $s(n)$ in (2) can be represented as a special AR(2p) model with zero input

$$\sum_{i=0}^{2p} a(i)s(n-i) = 0 \quad (13)$$

where $a(0) = 1$, and the polynomial

$$A(z) = \sum_{m=0}^{2p} a(m)z^{-m} \quad (14)$$

has complex conjugate roots at $z = e^{\pm j\omega_i}$ ($i = 1, 2, \dots, p$).

From [4], the correlation of the filtered output process $\tilde{y}(n)$ and the AR parameters in (13) satisfy

$$\sum_{i=0}^{2p} a(i)r_{\tilde{y}}(l-i) = 0 \quad l > 2p + n_d \quad (15)$$

Then from (15) the AR order $2p$ and AR parameters $a(i)$, $i = 1, \dots, 2p$ can be estimated by using SVD-TLS method, after which the harmonic frequencies can be estimated directly.

The following is our new prefiltering approach for harmonic retrieval in non-Gaussian ARMA noise

- 1) Obtain $\tilde{y}(n)$ from $y(n)$ using Hilbert transform.
- 2) Estimate $c_{4\tilde{y}}$ according to definition 1, then determine n_b and $b(i)$, $i = 1, \dots, n_b$ from (9) via SVD-TLS method.
- 3) obtain $\tilde{y}(n)$ from (10), Compute its correlation then determine the coefficients $a(i)$, $i = 1, \dots, 2p$ from (15).
- 4) Find roots of $A(z) = 0$ in (14), then compute frequencies of harmonics.

Numerical Examples

Simulation results are presented to illustrate the proposed algorithm in this section. We focus on demonstrating the performance of the new method when the colored non-Gaussian noise is symmetrically distributed or the harmonics are quadratic phase coupled. As most reported studies, we consider a time series of the form

$$y(n) = \sin(2\pi f_1 n + \varphi_1) + \sin(2\pi f_2 n + \varphi_2) + w(n) \quad (16)$$

which include two sinusoids. This makes only self phase/frequency coupling possible.

The noise $w(n)$ is an ARMA process with AR parameters $[1 \ -1.5 \ 0.8]$ and MA parameters $[1 \ -0.75 \ -2.5]$, its spectrum shows a sharp peak. $w(n)$ is asymmetrically distributed in example 1 and symmetrically distributed in example 2.

Example 1:

$f_1 = 0.23$, $f_2 = 2f_1 = 0.46$, $\varphi_1 = \pi/6$, $\varphi_2 = 0$, the harmonics are quadratic frequency coupled, $w(n)$ is exponential distributed, SNR=-3dB, parameter estimation results of the new prefiltering method are summarize in Table I.

Example 2: $f_1 = 0.26$, $f_2 = 0.3$, $w(n)$ is mixed Gaussian distributed. The detail about mixed Gaussian distribution can be found in [6], SNR=-5dB, the results of the new prefiltering method are given in Table II.

When the non-Gaussian noise is symmetrically distributed or the harmonics are quadratic phase(frequency) coupled, the noise AR parameter estimates are heavily biased if using the prefiltering method proposed in [4], which results the frequency estimates are often unobtainable. At the same time the new prefiltering method can obtain satisfied results.

Noise Model	b(1)	b(2)
True	-1.5	0.8
Estimate	-1.5617(0.1457)	0.8722(0.0693)
Frequency	f_1	f_2
True	0.23	0.46
Estimate	0.2306(0.0008)	0.4603(0.0009)

TABLE I STATISTICS OF AR PARAMETER AND FREQUENCY ESTIMATES OBTAINED VIA THE NEW PREFILTERING APPROACH(N=2048 IN EACH RUN, 20 MONTE-CARLO RUNS) IN THE CASE OF QUADRATIC FREQUENCY COUPLED HARMONICS.

Noise Model	b(1)	b(2)
True	-1.5	0.8
Estimate	-1.4773(0.1316)	0.8338(0.0929)
Frequency	f_1	f_2
True	0.26	0.3
Estimate	0.2601(0.0005)	0.3002(0.0006)

TABLE II STATISTICS OF AR PARAMETER AND FREQUENCY ESTIMATES OBTAINED VIA THE NEW PREFILTERING APPROACH(N=2048 IN EACH RUN, 20 MONTE-CARLO RUNS) IN THE CASE OF SYMMETRICALLY-DISTRIBUTED NON-GAUSSIAN NOISE.

Conclusion

In this paper a new approach is proposed to retrieve harmonics in colored linear non-Gaussian noise especially when the noise is symmetrically distributed or the harmonics are quadratic phase/frequency coupled. In order to estimate the non-Gaussian noise characteristics in the presence of harmonics, Hilbert transform is used to transform the real observations to their complex form. Then the higher order Yule-Walker equation of this complex process is established, which is used in the proposed prefiltering method to retrieve harmonics. Simulation results have shown the effectiveness of the new method.

Appendix

The k th-order cumulant of a complex process has 2^k different definitions depending on the position of conjugations, we now prove that none of them satisfies the Yule-Walker equation in (9)

Proof: The k th-order cumulant of the complex process $\tilde{w}(n)$ in (6) is defined as

$$C_{k,\tilde{w}}(\tau_1, \dots, \tau_{k-1}) = Cum(\tilde{w}^{(p_0)}(l), \dots, \tilde{w}^{(p_{k-1})}(l + \tau_{k-1}))$$

where $p_i \in \{1, -1\}$, $i = 0, \dots, k-1$.

$$\tilde{w}^{(1)}(t) = \tilde{w}(t), \tilde{w}^{(-1)}(t) = \tilde{w}^*(t)$$

then $C_{k,\tilde{w}}(\tau_1, \dots, \tau_{k-1})$

$$= \sum_{m_1} \dots \sum_{m_{k-1}} C_{k,w}(\tau_1 - m_1, \dots, \tau_{k-1} - m_{k-1}) \cdot C_{k,h_1}(m_1, m_2, \dots, m_{k-1}) \quad (A1)$$

where $C_{k,h_1}(m_1, m_2, \dots, m_{k-1}) =$ (A2)

$$\sum_{j_0} h_1^{(p_0)}(j_0) h_1^{(p_1)}(j_0 + m_1) \dots h_1^{(p_{k-1})}(j_0 + m_{k-1})$$

from[2], $C_{k,w}(\tau_1, \dots, \tau_{k-1})$

$$= r_{k,e} \sum_{n=0}^{\infty} h_w(n) h_w(n + \tau_1) \dots h_w(n + \tau_{k-1}) \quad (A3)$$

where $r_{k,e} = E[e(n)^k]$

then $\sum_{i=0}^{n_b} b(i) C_{k,\tilde{w}}(\tau_1 - i, \tau_2, \dots, \tau_{k-1}) =$

$$\sum_{m_1} \dots \sum_{m_{k-1}} r_{k,e} \sum_{n=0}^{\infty} h_w(n) d(n + \tau_1 - m_1) \dots h_w(n + \tau_{k-1} - m_{k-1}) \cdot C_{k,h_1}(m_1, m_2, \dots, m_{k-1}) \quad (A4)$$

as we all know, $d(i) = 0$ when $i > n_d$ or $i < 0$, but it is obvious that the last equation is not identical with zero no matter what the τ_1 is. From (A1), the way of cumulant definition has relation with only C_{k,h_1} , so none of these definitions satisfies the Yule-Walker equation.

Reference

- [1] E. J. Wegman, S. C. Schwartz and J. B. Thomas, Topics in Non-Gaussian Signal Processing. New York : Springer-Verlag, 1989.
- [2] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some application", Proc. IEEE, vol. 79, pp. 278-305, 1991.
- [3] B. Giannakis and J. M. Mendel, "Cumulant-based order determination of non-Gaussian ARMA model", IEEE Trans. Acoustic Speech and Signal Processing, vol. 38, pp. 1411-1422, 1990.
- [4] X.-D. Zhang, Y.-C. Liang and Y.-D. Li, "A Hybrid Approach To Harmonic Retrieval In Non-Gaussian ARMA Noise," IEEE Trans. Inform. Theory, Vol. 40, pp. 1220-1226, 1994.
- [5] B. Sadler, G. Giannakis and S. Shamsunder, "Noise Subspace Techniques in Non-Gaussian Noise Using Cumulants," IEEE trans. Aerospace and Electronic Systems, vol. 31, pp. 1009-1018, 1995.
- [6] B. Sadler, "Detection in Correlated Impulsive Noise using Fourth-Order Cumulants," IEEE Trans. Signal Processing, vol. 44, pp. 2793-2800, 1996.

The Use of Arbitrary Sequences for Real-Time Wireless Channel Evaluation

A Al-Dabbagh and M Darnell

The Institute of Integrated Information Systems
School of Electronic and Electrical Engineering
University of Leeds, Leeds
LS2 - 9JT, UK

tel: (44) -113-233 2075

fax: (44) -113-233 2032

Email : eenaa@electeng.leeds.ac.uk

Abstract

The need for accurate channel measurements for broadband communication system is well known. The problem of system identification constitutes the basis for any wireless channel measurement and estimation. The classical approach to system identification employs a pseudo-random (PR) sequence as a test signal. The system response to this PR sequence is then correlated with a replica of the test signal in order to obtain an estimate of the system impulse response. The fundamental assumption upon which this technique is based is that the PR test signal has an ideally impulsive autocorrelation function (ACF). In practice, the majority of sequences of interest do not possess this idealised ACF. The consequences of employing test signals with non-impulsive ACF results in the introduction of estimation errors. The problem is further accentuated when the classical approach to system identification is applied to channel estimation. In this paper, the problem of system identification is analysed for the case when the ACF of the test signal is non-ideal. It is shown that the system impulse response estimation error can be reduced by using a compensation filter whose response is has been related to the imperfection in the test signal ACF. The applicability of this approach to broadband channel estimation is then demonstrated.

Key words : Channel Evaluation, System Identification, Correlation

Introduction

Radio data transmission without knowledge of the communication channel parameters is difficult and will most certainly lead to low-reliability communications. In the case of digital information transmission, the problem is further accentuated, especially when higher data rates are sought [1]. The limitation imposed on the reliability and throughput of a communication process by the channel has led to an increasing adoption of channel evaluation techniques as part of the overall communication system design and development [2].

Many receiver designs which are intended to operate over bandlimited and perhaps time-varying, channels require an estimate of the prevailing channel conditions; these are completely characterised by the scattering function of the channel. The bandlimited type of channel can be influenced by the presence of multipath components in the received signal, which result in the introduction of inter-symbol-interference (ISI). For example, in the case of the

Rake receiver implementation, accurate channel measurements are needed [3] to improve performance over channels with ISI. In the case of time-varying channels, the channel characterisation must be performed in real time in order to maintain an up-to-date estimate of the channel; the system may then be adapted in response to such measurements.

The problem of linear time-invariant (LTI) system parameter identification forms the basis of wireless channel evaluation. Classical parameter identification assumes that the system-under-test is of a baseband nature, whereas wireless channels are a subclass of bandpass systems and hence their unknown real impulse response at radio frequencies can be represented by an equivalent complex baseband impulse response. This difference between baseband and bandpass system identification has important implications for the migration of baseband system identification methods to radio channel evaluation.

The classical approach to system identification typically employs some form of pseudo-random (PR) sequence as a test signal. The system response to this PR sequence is correlated with a replica of the test signal in order to obtain an estimate of the system impulse response. The fundamental assumption upon which this technique is based is that the PR test signal has an ideal autocorrelation function (ACF). In practice, the majority of sequences of interest do not possess this ideal ACF, and the consequences of employing test signals with non-impulsive ACF involve in the introduction of estimation errors. The problem is further accentuated when the classical approach to system identification is applied to evaluation of radio channels.

System Identification

Let $h(t)$ and $H(f)$ be the corresponding unit impulse and frequency domain responses of an unknown linear time-invariant (LTI) system, respectively, related by the Fourier analysis. Linear system theory shows that if the input ACF is a perfect Dirac delta function, then the crosscorrelation between the system input, $x(t)$, and the corresponding output, $y(t)$, is

$$R_{xy}(t) = \int_{T_p} h(u) R(t+u) du \quad (1)$$

where $R_{xx}(\tau)$ is the ACF of the period ie test signal $x(t)$ with period T_p . The variable u is a dummy time variable and τ is a delay variable. The estimate of the impulse response, $\bar{h}(t)$, will be directly proportional to $h(t)$ if $R_{xx}(\tau)$ is perfectly impulsive. $R_{xx}(\tau)$ is periodic if $x(t)$ is periodic and, in general, can be expressed as

$$R_{xx}(\tau) = \delta(\tau) + g(\tau) \quad (2)$$

ie, as a sum of a unit impulse and a sidelobe function $g(\tau)$ which is zero at $\tau=0$. Substituting 2 into 1 yields

$$\bar{h}(\tau) = h(\tau) + \int_{T_p} h(u) g(\tau+u) du \quad (3)$$

Eq 3 shows that the $\bar{h}(t)$ is made up of the required response $h(t)$ and an error term, $\int_{T_p} h(u) g(\tau+u) du$. If

this error term is reduced to zero, perfect impulse response estimation is achieved. Now applying $\bar{h}(t)$ to a *compensating* filter with an impulse response $q(t) \leftrightarrow Q(f)$, it may be readily shown that the output of this filter, $z(t)$, is given by

$$z(t) = h \otimes \underbrace{[q + g \otimes q]}_{\psi(t)} \quad (4)$$

For $z(t)$ to equal $h(t)$, one requires that $\psi(t) = \delta(t)$. By substituting for $\psi(t)$ taking Fourier transforms and re-arranging, it can be shown that perfect system identification is possible using the system configuration of figure 1 where the impulse response of the post-filter may be shown to be

$$Q(f) = \text{invFourier} \left[\frac{1}{1 + G(f)} \right] \quad (5)$$

This result shows that any periodic test signal with an imperfect periodic ACF can be used to give accurate estimates of system impulse response, providing that eq 6 has no singularities, ie $G(f) \neq -1$, for all f .

Simulation Results

A 7-level m-sequence of length 342 digits was used in the identification of an LTI system with the impulse response $h(t) = t e^{-0.04t}$. Its periodic ACF may be expressed by

$$R_{xx}(\tau) = \begin{cases} 1 & \text{when } \tau = 0 \\ -0.07143 & \text{when } \tau = \pm 57 \\ +0.07143 & \text{when } \tau = \pm 114 \end{cases} \quad (6)$$

The estimated impulse responses due to the classical and new methods are shown in figure 2 along with the theoretical response. The results clearly show the improvement in the estimation accuracy of the impulse response of the system under test.

This form of compensated system identification, in which the error in the estimated impulse response is expressed in terms of the imperfection in the ACF of the test signal has been published previously by the researcher in [4]. It was further developed and reformulated in terms of the concept of inverse filtering in [5] where it was termed *generalised system identification*.

Channel Evaluation

Broadband communication channels are often modelled as a linear system which is fully characterised by its impulse response. Channel evaluation is the process of estimating the channel impulse response. Normally, this is carried out at baseband. Since the impulse response is a real function at RF, the underlying theory of system identification, which is carried out at baseband, can

be applied successfully to channel estimation. The results may then be expressed in terms of the channel baseband equivalent impulse response which is a complex function. This approach has been pursued by the researchers and the concept of *generalised channel evaluation* is being introduced.

As an example, consider the test waveforms in figure 3a and 3b which are applied to the channel inphase and quadrature inputs respectively; these test signals do not possess an impulsive ACF. It is, however, possible to obtain accurate channel measurements using these imperfect waveforms using generalised channel evaluation. The evaluation of a simulated arbitrary radio channel was carried out using the two waveforms and the results are shown in figure 4a and 4b for the channel inphase and quadrature impulse responses respectively.

Final Remarks

In this paper, we have demonstrated the use of test signals with non-impulsive ACF to obtain an accurate estimate of the impulse response of an unknown baseband system. The same concept was then applied to the problem of channel evaluation, and it was shown that it is possible to obtain accurate estimate of the channel impulse response using probing signals with arbitrary ACFs.

References

- [1] Proakis, J : "*Digital Communications*", McGraw-Hill International Editions, 3rd Edition, 1995.
- [2] Darnell, M : "*Real-Time Channel Evaluation*", Modern HF Communications, AGARD Lecture Series LS-127, NATO, 1983.
- [3] Griparis, T : "*Channel Estimation and Tracking Algorithms for DS-CDMA Communication Systems*", Masters Thesis, King's College, London, England, 1996-1997.
- [4] Al-Dabbagh, A and Darnell, M and Noble, A and Farquhar, S : "*Accurate System Identification using Inputs with Imperfect Autocorrelation Properties*", Electronic Letters, Aug., vol. 33, no. 17, pp. 1450-1451, 1997.
- [5] Al-Dabbagh, A and Darnell : "*The Theory and Application of Reciprocal Pairs of Periodic Sequences*", Cryptography and Coding VI, Springer, ISBN: 3-540-63927-6", 1997.
- [6] Lee, Y W : "*Statistical Theory of Communication*", Wiley, 1960.

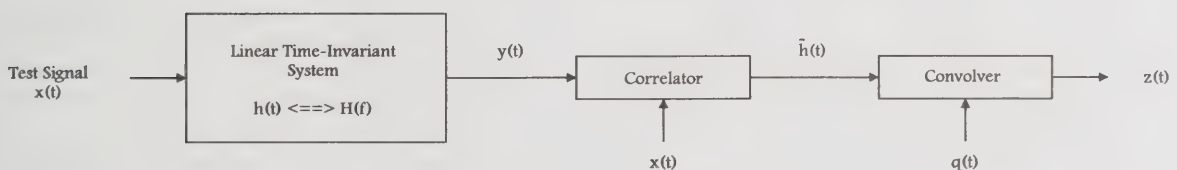


Fig 1 - Compensated System Identification.

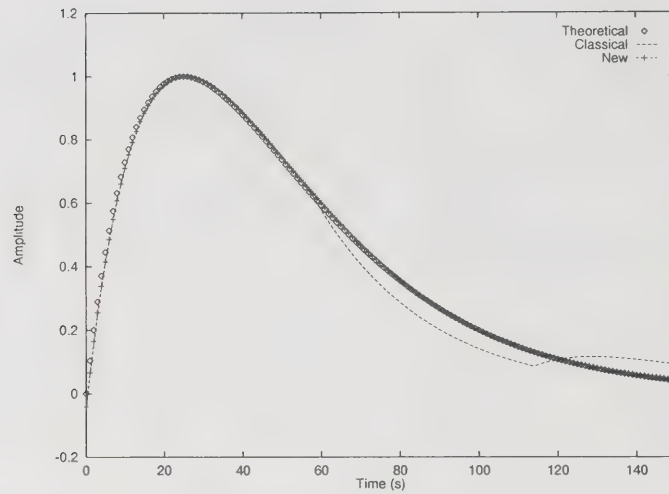


Fig 2 - System Identification Compensated Methods.

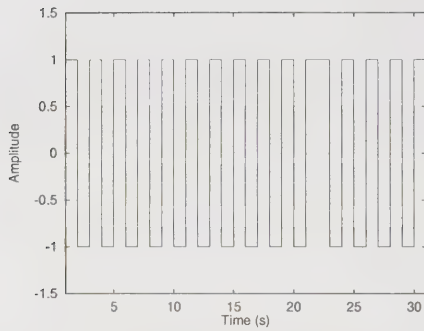


Fig 3(a) The Inphase Test Signal.

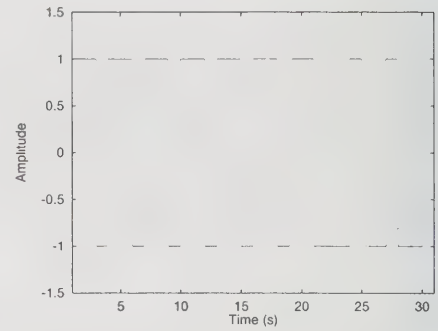


Fig 3(b) The Quadrature Test Signal.

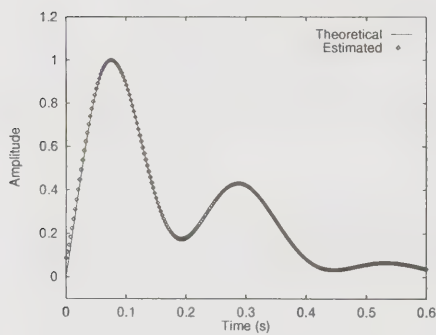


Fig 4(a) The Inphase Channel Impulse Response.

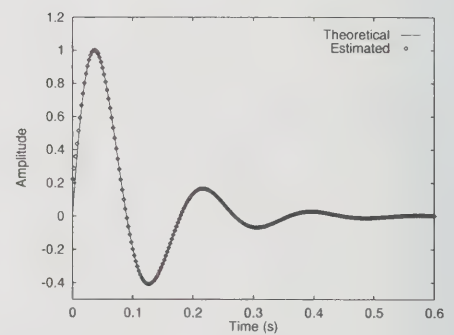


Fig 4(b) The Quadrature Channel Impulse Response.

Frequency-Domain Delay Distortion Equalization - The Complex Spectrum Regeneration Method

Masami Akaike

Dept. of Elec. Eng., Science University of Tokyo,

1-3 Kagurazaka, Shinjuku, Tokyo 162-8601 JAPAN

Tel: +81 3 3260 4272 Ext. 3328 · Fax: +81 3 3260 4607 E-mail:akaike@ee.kagu.sut.ac.jp

ABSTRACT

In radio communications, equalization of delay distortion due to multipath propagation is one of the most necessary and difficult problems. There are two equalization methods, i. e., the time-domain method and frequency-domain method. In this paper, the frequency-domain linear equalization, in which equalization is performed by deconvolution of the received signal by the transfer function of the propagation path is presented. In the frequency domain, deconvolution is done by division by the transfer function. Since this method is made by software-based devices, faster, stabler, and more adaptive equalization is possible.

1. INTRODUCTION

In radio communication systems design, the propagation path is characterized by multipath propagation, in which the main wave is accompanied by delay waves. Multipath propagation causes delay distortion in transmission waveforms. In radio communications, especially in mobile communications, delay-distortion equalization is one of the most necessary and fundamental problems for obtaining high-quality transmission characteristics.

There has been the time-domain equalization in which equalization is performed symbol by symbol in the time sequence by use of a transversal filter. In this method (more specifically, linear time-domain method), equalization is made by the deconvolution by the impulse response of the propagation path. The time-domain equalization has a long history and many works have been presented[1].

Recently, the frequency-domain equalization has been attracting attention, since the equalization in the frequency domain is performed by digital signal processing by software-oriented devices and therefore has some merits compared with the time-domain method. There have been some works published so far[2][3]. From mathematical duality, deconvolution by the impulse response in the time domain is equivalent to division by the transfer function in the frequency domain.

It is the purpose of this paper is to propose and discuss a linear frequency-domain method, in which the complex spectrum of the main wave is regenerated by deconvolution of the received spectrum by the transfer function of multipath propagation. Since the deconvolution is a division by a complex transfer function in the frequency domain, and the division is easily fulfilled by digital computer operation, this equalization method offers a higher-speed, stabler and more adaptive equalization regardless of the values of amplitude and delay time. In this paper, the basic consideration of the frequency-domain method and the performance characteristics obtained by computer simulation are presented.

2. MULTIPATH PROPAGATION

Figure 1 shows a simplified multipath propagation model. The wave reaching the receiver (R) from the transmitter (T) consists of multiple waves, i. e., the main wave and delay waves, due to reflections and/or diffractions of the buildings, trees and the ground. Each wave has different delay time, which causes delay distortion in received symbols.

3. BASIC CONSIDERATION OF COMPLEX SPECTRUM REGENERATION

3.1 Comparison of the time-domain and frequency-domain methods

It is well known that the time-domain operation and the frequency-domain operation have a dual relationship and each operation is connected by the Fourier transform. Comparison of the time-domain and frequency-domain methods is shown in Table 1. In the time-domain method, the characteristics of the propagation path are expressed by the impulse response, and the equalization

(deconvolution) is approximated by the convolution of the inverse characteristics. In the frequency-domain method, on the other hand, the propagation characteristics are expressed by the transfer function. In the linear equalization, deconvolution is done by division by the transfer function. The complex spectrum having no distortion is regenerated from the distorted received spectrum divided by the transfer function of the propagation path. Equalization in the frequency domain is performed by digital computation by software-based devices.

3.2 Complex spectrum regeneration

Complex spectrum is regenerated under the following basis:

- (1) One frame consists of one training burst and more than one information bursts. The training burst consists of an already-known sequence of symbols.
- (2) Delay is equalized burst by burst (not symbol by symbol).
- (3) The training burst is used for estimating the transfer function of the propagation path. The transfer function is assumed to be unvarying for the training burst and information burst.
- (4) For realizing satisfactory equalization, guard bits longer than the delay time are embedded in the head and tail of a burst.
- (5) For the Fourier analysis of the waveform of the burst, a symbol is sampled by appropriate times.

Let $f_T(t)$, $g_T(t)$, $f(t)$, and $g(t)$ be the waveforms of the training burst, received burst for the training burst, information burst, and received burst for the information burst, respectively. Then they are expressed as

$$g_T(t) = f_T(t) + \rho f_T(t-\tau), \quad g(t) = f(t) + \rho f(t-\tau) \quad (1)$$

where ρ and τ are the relative amplitude and delay time of delay waves with respect to the main wave; $\rho f_T(t-\tau)$ and $\rho f(t-\tau)$ are representative expressions for multiple delay waves. From Eq.(1), one obtains corresponding Fourier series expression as

$$\sum_k G_{Tk} = \sum F_{Tk}(1 + \rho \exp(-jk\Omega_0\tau)), \quad \sum_k G_k = \sum F_k(1 + \rho \exp(-jk\Omega_0\tau)) \quad (2)$$

where \sum_k expresses the sum with respect to k ; F_{Tk} , G_{Tk} , F_k , and G_k are Fourier components of $f_T(t)$, $g_T(t)$, $f(t)$, and $g(t)$, respectively; the sum is taken with respect to k ; and Ω_0 is the fundamental angular frequency for one burst. From Eq.(2), one obtains

$$F_k = G_k F_{Tk} / G_{Tk} = G_k / (G_{Tk} / F_{Tk}). \quad (3)$$

In Eq.(3), G_{Tk} / F_{Tk} is the transfer function of the propagation path. From Eq.(3), the spectrum of the information burst, F_k , having no delay distortion is obtained from received waveform spectrum, G_k , divided by the transfer function, G_{Tk} / F_{Tk} . The delay-equalized waveform is calculated by the inverse Fourier transform of F_k .

The transfer function G_{Tk} / F_{Tk} can be expressed by two forms, i. e., (i) a series of complex number from direct calculation of G_{Tk} / F_{Tk} and (ii) a form of a function, $(1 + \sum_i \rho_i \exp(-jk\Omega_0\tau_i))$, properly estimated from measured G_{Tk} / F_{Tk} .

4. COMPUTER SIMULATION

The block diagram of the simulation is shown in Fig. 2. The transmission burst is made by a pulse generator (PUL) driven by a binary (0 or 1) random-number generator (RND) in which the probabilities of generation of 0 and 1 are equal. A pulse is expressed by a sinc function for symbol 1:

$$\text{sinc}(t) = \sin \pi t / (\pi t).$$

The transmission burst is divided into two branches, one of which forms the main burst (MBUS) and the other is delayed bursts (DBUS). The bursts from two branches are added and fed to the receiver and equalized. The output of the transmitter and the input of the receiver are represented by T and R, respectively (see Fig.1). Therefore, the block T to R represents the propagation path. A white noise is added at the input (point R) of the receiver. The white noise is expressed as binary symbols of random-sequence uncorrelated to the main burst.

In the equalizer block (FEQL), the Fourier transform and calculation of Eq.(3) are performed. The waveforms and eye apertures at T, R, and S are observed.

5. SIMULATION RESULTS

Figure 3 shows the waveforms and eye apertures at T, R, and S, for the case where there are three

delay waves, i. e., $\tau_1/T=3.3$, $\tau_2/T=-4.4$, $\tau_3/T=6.3$ (T = one symbol time-length), and $\rho_1 = \rho_2 = \rho_3 = 0.999$. This is relatively a severe case where delay times are long and the amplitudes of delay waves are close to the main-wave amplitude. The length of one burst is of 64 symbols and one symbol is sampled twice. As an ideal case, the level of the white noise is zero.

Comparing three waveforms and three eye apertures, one can see satisfactory equalization obtained. Since the equalization process of this method is a simple division of complex number, good equalization is always possible regardless of the number of delay waves, their amplitudes and delay times. In frequency-domain operation there are no problems even when the amplitude of delay wave is larger than that of the main wave, unlike in the time-domain operation. Since the equalization is performed only by calculation of complex number, this equalization offers a higher-speed, stabler and more adaptive equalization regardless of the values of amplitude and delay time.

Eye apertures when a white noise is added to the received burst at R are shown in Fig.4. The delay waves condition is the same with Fig.3. Figure 4 shows two cases. Eye aperture [A] shows the case where the obtained transfer function is assumed to exclude the noise by some nonlinear operation. Eye aperture [B] shows the case where the transfer function is directly calculated from G_{Tk}/F_{Tk} . In this case the noise is included in the obtained transfer function. One understands that only the linear operation degrades the equalization due to noise when the multipath condition is very severe. However, when the amplitudes of the main wave and delay waves are not so close, the linear operation in the frequency domain gives sufficiently satisfactory equalization.

6. CONCLUSION

For the frequency-domain delay-distortion equalization, a method by regenerating complex spectrum is proposed and fundamental performance characteristics are discussed. Since the deconvolution of the propagation path characteristics is realized by the division by the transfer function, faster, stabler and more adaptive equalization is expected compared with the time-domain method using transversal filters.

REFERENCES

- [1] Shahid U. H. Qureshi, "Adaptive equalization", Proc. IEEE, Vol.39, No.9, pp1340-1387, Sept., 1985.
- [2] T. Nakazato, H. Ochi, and S. Kinjo, "A new adaptive equalizer realizing desired frequency response," Proc. 1996 Comm. Soc. Conf. of IEICE, A-118, Sept., 1996.
- [3] H. Ohta, et. al., "Mobile reception of digital terrestrial television broadcasting," Technical Report of IEICE, RCS97-78 (1997-7)

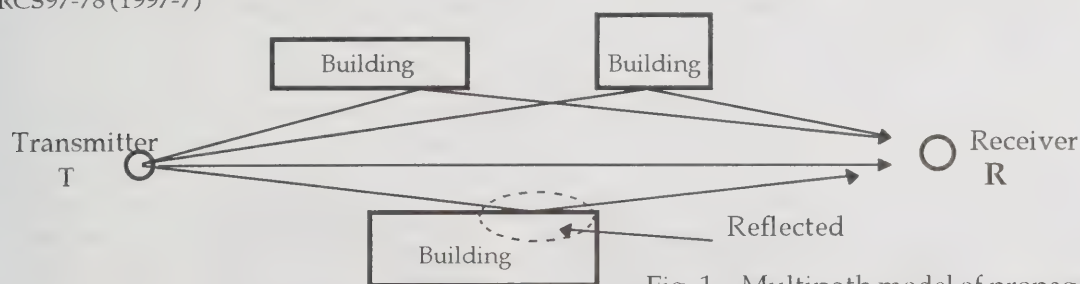


Fig. 1 Multipath model of propagation path.

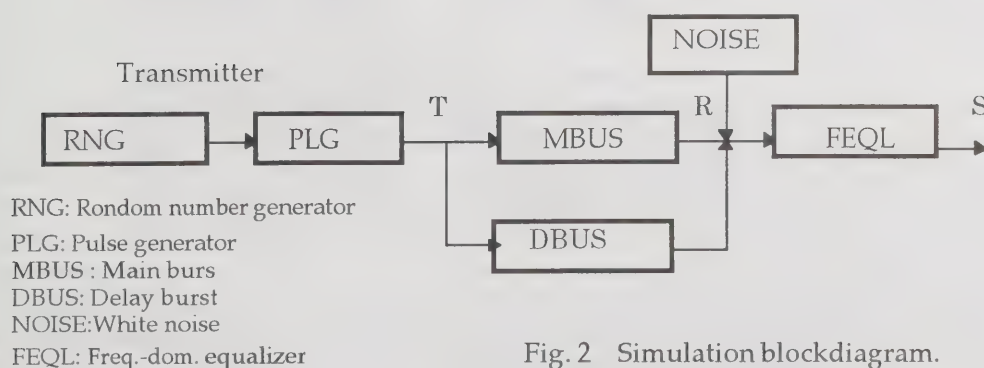


Fig. 2 Simulation blockdiagram.

Table 1 Comparison of time-domain and frequency-domain methods

	Time domain	Frequency domain
Function to be equalized	Waveform (function of time)	Complex spectrum (function of frequency)
Propagation path expression	Impulse response	Transfer function
Equalization* (Deconvolution*)	Convolution of inverse characteristics	Division
Equalizer device	Transversal filter (hardware based)	Computation (software based)

(*)Note: Linear equalization is considered here.

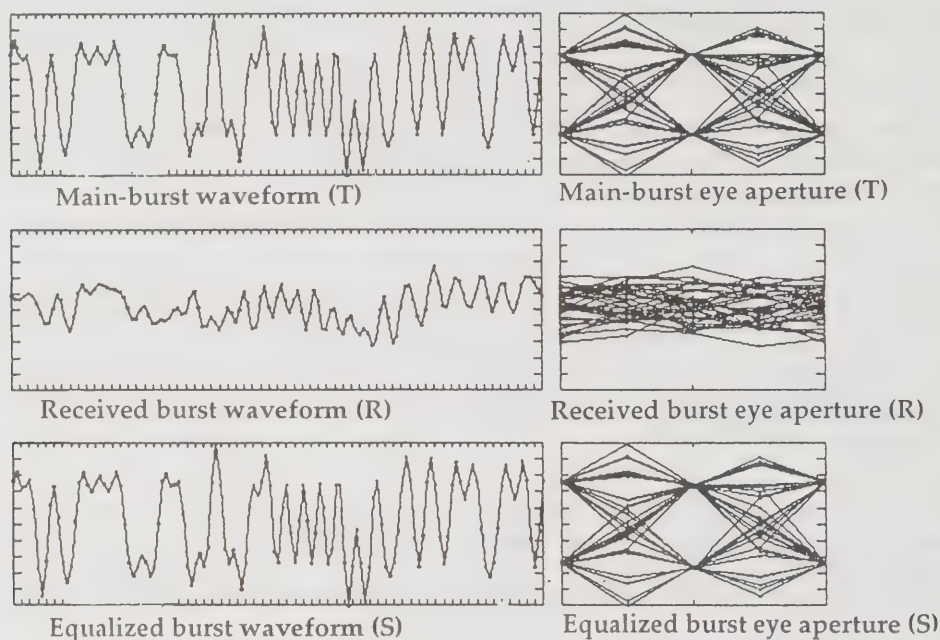


Fig. 3 Waveforms and eye apertures for $\tau_1/T=3.3$, $\tau_2/T=-4.4$, $\tau_3/T=6.3$, $\rho_1=\rho_2=\rho_3=0.9$

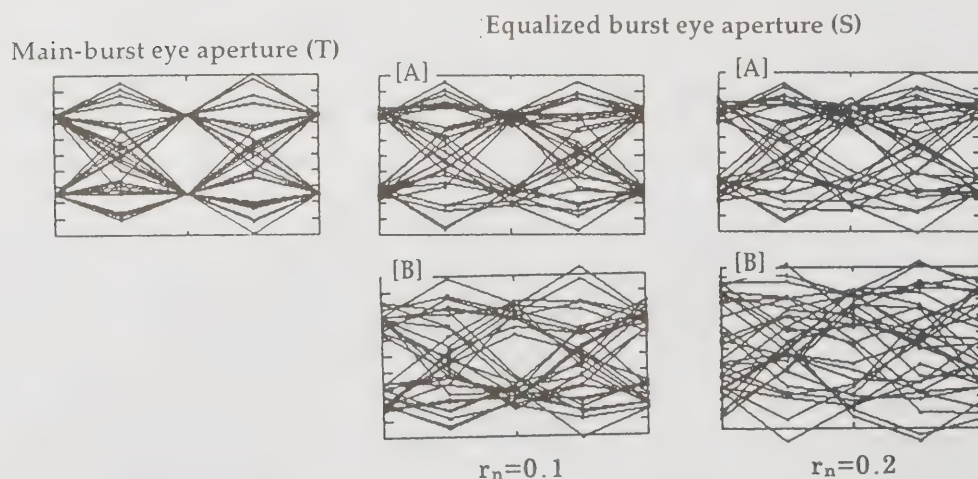


Fig. 4 Eye apertures with white noise ($\tau_1/T=3.3$, $\tau_2/T=-4.4$, $\tau_3/T=6.3$, $\rho_1=\rho_2=\rho_3=0.999$). [A]: White noise is eliminated from the obtained transfer function. [B]: White noise is included in the obtained transfer function. (r_n =relative amplitude of the white noise with respect to the main burst)

Frustum Ray Tracing Technique for High Spatial Resolution Channel Characteristic Map

Hajime Suzuki and Ananda S. Mohan*

*Telecommunications Group, Faculty of Engineering
University of Technology, Sydney*

Abstract— *A novel frustum ray tracing technique is introduced that utilizes fast line clipping algorithm rather than conventional time-consuming ray intersection test algorithm. The usefulness of the technique in predicting channel characteristic maps with high spatial (one fifth of wavelength) resolution incorporating hundreds of thousands receiving points in a complex in-building environment is demonstrated by results predicted by a software program that runs on a typical PC.*

I. INTRODUCTION

High data speed and high user density in-building radio communication systems have been recognized as near-future services that provide users with wireless accesses to the various networks. One of the major challenges in archiving such systems is to accurately characterize the radio propagation channels in the complex in-building environment where various reflecting and obstructing objects exist between the transmitter and the receiver. The ray tracing technique has been widely adopted to predict the static narrowband/wideband channel characteristics of the in-building radio channels. The role of the ray tracing algorithm is to efficiently find valid geometric optics (GO) propagation path between the transmitter and the receiver. Over the past years, improvements of the technique have been reported by several researchers enhancing the efficiency and accuracy of the prediction while accommodating a larger number of receiver locations and more complex environment.

Without any enhancement technique, the computational complexity of an image source based ray tracing technique can be expressed as $y \times n^r$, where y is the number of receiving points and n is the number of objects in the model, which can become unmanageably high when applied to a complex building model. In addition, its calculation time is linearly proportional to the number of receiving points, thus making it unsuitable when a large number of receiving points are involved. Although, a computational acceleration technique was introduced in [1] by incorporating two grids of different resolutions to define receiving points, a finer grid is required in order to guarantee that no reflecting object is omitted from the

calculation, which in turn does not contribute to the prediction efficiency.

The computational burden of algorithms such as ray launching technique [2], shooting and bouncing ray with image method [3], and ray tube tracing technique [4] that cast a limited number of rays or ray tubes separated by semi-uniform angles can be expressed as $x \times 2^r$ where x is the number of rays or ray tubes launched from the source and r is the number of multiple reflections and/or transmissions. Although, the complexity of the building model and the number of receiving points have lesser impact on the computational load in this case, the techniques suffer from the errors caused by finite ray separation angles as shown in Fig. 1 (a).

Figure 1 (a) shows a received signal power map at 1 GHz predicted by the ray tube tracing technique [4] on a modeled floor of Level 23 in the University of Technology, Sydney's Tower Building. The floor is called the target environment in this paper. The received power is normalized by the power at 1 m free space path. The floor model consists of 498 patches and a detailed description of construction materials is presented in Section III. A vertically polarized dipole antenna transmitter is placed at $x = 12.1$ m and $y = 8.6$ m while its height is 1.6 m from the floor. 20,480 ray tubes were defined and cast from the transmitter. Only transmitted signals through walls/separations were modeled in the prediction results shown in Fig. 1 in order to identify and clarify the problem. The receiving points are defined at a height of 1.6 m from the floor with a resolution of 0.2 wavelength (6 cm) resulting in 326,041 receiving points. Calculation is performed using MATLABTM running on a 200 MHz

*Dr. Ananda S. Mohan. Phone: +61-2-9514-2447. Fax: +61-2-9514-2435. E-mail: ananda@eng.uts.edu.au.

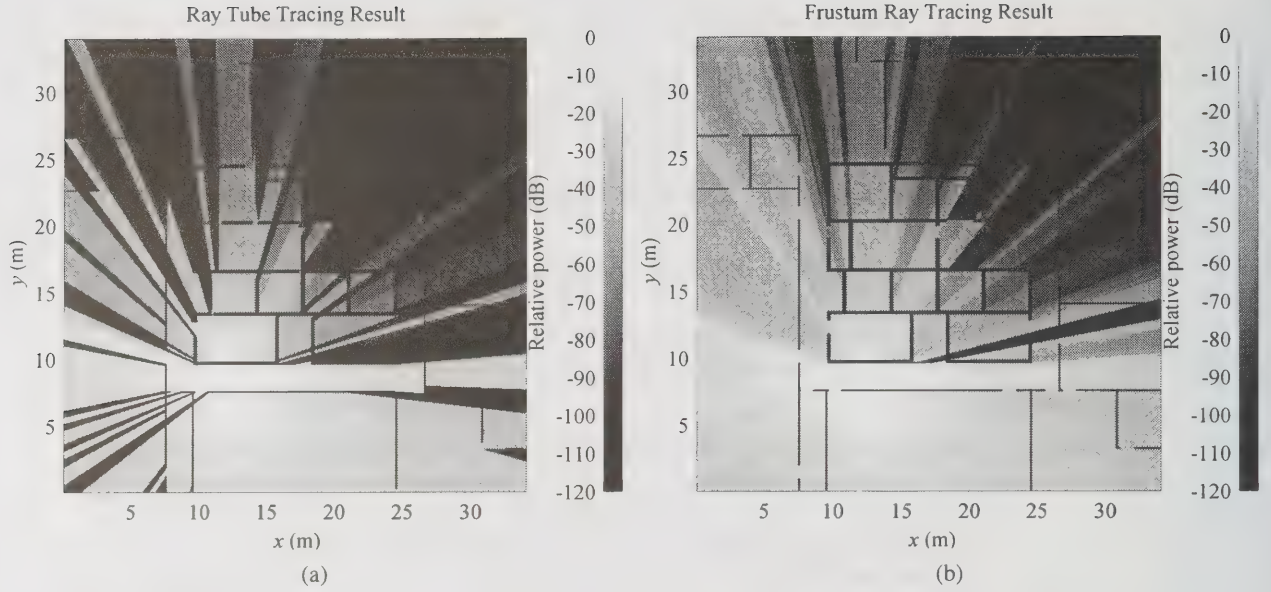


Figure 1: Comparison on the prediction of direct received signal power by ray tube tracing method and frustum ray tracing method.

Pentium™ PC and the calculation time is around 2 hours. When a ray tube is traced farther from the location of the base station, its wave front becomes larger and when it interacts with any wedge or a corner in the model, the tracing of the ray tube terminates, leaving regions that cannot be reached by any ray tube, as indicated by the black areas in Fig. 1 (a). Receivers located in these regions are predicted to receive no signal at all, thus resulting in an erroneous prediction. This problem is common in any ray tracing technique that relies on the finite discrete separation angles of rays or ray tubes. The problem can be remedied by increasing the number of initial rays, however it directly increases the computational burden. The rays or ray tubes can be extended, instead of being terminated, at a wedge or corner to reach every area of the floor, however, in this case, signals that should be blocked otherwise could appear as if no blockage exists, causing prediction errors. In order to circumvent this problem, we have developed the new frustum ray tracing technique for efficient prediction of channel characteristic maps.

II. FRUSTUM RAY TRACING TECHNIQUE

Figure 2 (a) shows an example environment of a single rectangular room which contains a wall. The frustum ray tracing technique (FRTT) algorithm starts from enclosing the entire model space by a rectangular box with six faces. In our case they coincide with the six faces of the room. Then the model space is split into six pyramids each of whose apexes coincide with the location of the source antenna and whose base face coincides with one of the six faces of the model-enclosing-box as shown in Fig. 2 (a). The problem is to find any edge, wedge, or corner formed

by the modeled objects inside the pyramids. This is performed by applying the Cyrus-Beck clipping algorithm [5] to every edge in the model space to be clipped by the pyramid. In Fig. 2 (b), the edge $L-M$, which is a part of the wall $L-M-N-O$, is clipped by the pyramid $Tx-A-B-C-D$ resulting in the edge $H-I$ illuminated by the source. If any edge, wedge, or corner is found to lie inside the pyramid, the pyramid is split into two pyramids by the plane defined by the source point and two end points of the edge, which are indicated by Tx , H and I in Fig. 2 (b). Two new pyramids are created; the first is pyramid $Tx-A-B-F-E$ and the second is pyramid $Tx-E-F-C-D$ as shown in Fig. ?? (b). The pyramid $Tx-E-F-C-D$ contains a reflecting face $H-I-J-K$. When no edge, wedge, or corner is found inside the pyramid, the electric field is calculated at every receiving point inside the pyramid. The receiving points inside the pyramid $Tx-A-B-F-E$ and $Tx-H-I-J-K$ are not obstructed and hence represent an LOS case. When a pyramid is intersected by a plane, a frustum is created. Every receiving point inside the frustum $H-I-J-K-E-F-C-D$ is obstructed by the wall thus representing a non-LOS case. The receiving points inside the pyramid or frustum can be easily found by enclosing the model space by four or more planes that coincide with the faces of the pyramid or frustum. After determining received electric field inside the pyramid, the first order image source Tx' is created by reflecting the source point Tx with respect to the wall, as shown in Fig. 2 (c). The base of the pyramid $Tx'-H-I-J-K$ is extended so that it lies on the plane coincides with one of the faces of the box enclosing the model as shown in the Fig. 2 (c). Edges, wedges, or corners inside the frustum $H-I-J-K-P-Q-R-S$ are then sought for.

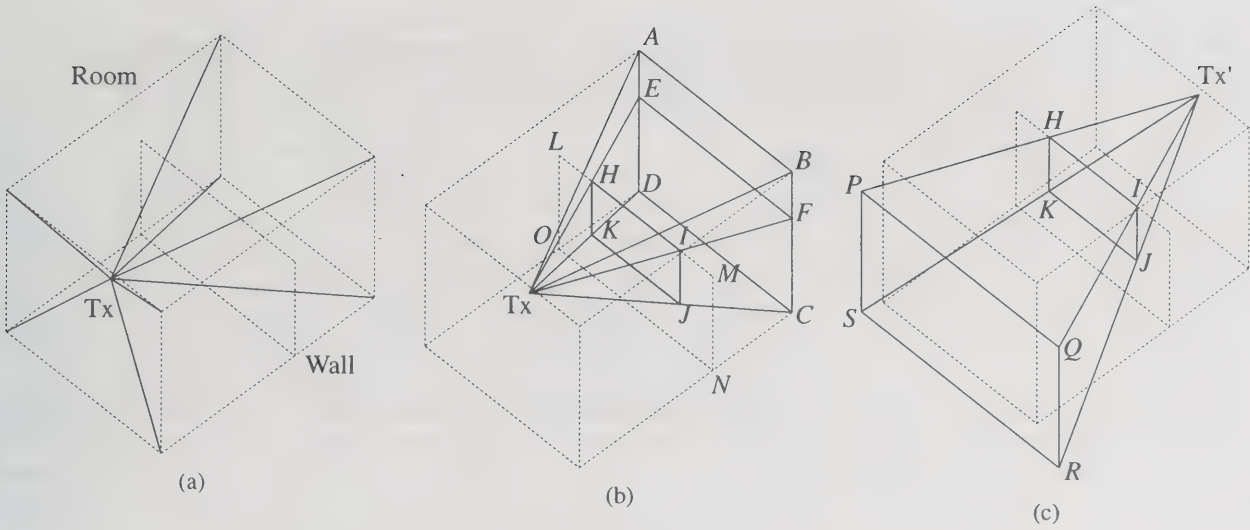


Figure 2: Pictorial presentation of frustum ray tracing algorithm.

This process is performed until an allowed maximum tracing depth is reached or the attenuation of the signal associated with the frustum reaches a preset maximum value.

Unlike the ray tracing techniques, that basically trace rays or ray tubes, the new technique traces pyramids or frustums, and thus is named as Frustum Ray Tracing Technique. The main advantages of FRTT include its accuracy as it predicts the GO path exactly and its capability of guaranteeing that no receiving point or modeled object is missed. Also FRTT is more efficient than the conventional ray tracing techniques because it does not rely on the intersection test of rays which usually consumes major portion of calculation time in a ray tracing simulation. However, the calculation time of the FRTT grows exponentially with the number of multiple-reflections, thus making the technique most suitable when smaller number of multiple-reflections are involved. In this paper, we consider a case in which the maximum order of multiple-reflections is four which is verified by the comparison with the measurement results as discussed in the following section.

Figure 1 (b) shows the result predicted by FRTT performed on the target environment. The computation time is less than 1 hour, which is a saving of 50% of the computation time needed for ray tube tracing prediction. It can be seen that the transmitted signals reach every receiving point on the entire floor (except the area obstructed by conductors) as expected from the GO solution. Since it is a full-three-dimensional technique, FRTT can easily be used even for predictions involving multiple floors.

III. COMPARISON

In order to verify the applicability of the FRTT to the real environment, narrowband received signal power measurements were conducted on the actual floor of the target en-

vironment and the results are compared with the prediction. The measurement area was chosen to include LOS, non-LOS, and around corner as shown in Fig. 3. Note that Fig. 3 is a section of a floor plan shown in Fig. 1, and the x - y axis is consistent between the two figures. The received signal power was measured by a HP 8720A vector network analyzer and computer-controlled antenna scanner. Measurement was performed during the night to ensure the minimum interference with the movement of people and/or equipment.

Prediction is performed on the modeled floor and up to four multiple reflection and any number of multiple transmission is incorporated. The major structures of the floor consist of 10 cm thick brick wall and 25 cm thick reinforced concrete wall. Floors and ceiling are made of 50 cm thick reinforced concrete. Small objects such as wooden doors and aluminum frames are included in the simulation. The objects are modeled as homogeneous dielectric slabs assigned with appropriate dielectric parameters. Figure 4 shows a comparison of measured and predicted received signal power map. The fast fluctuation of small-scale (less than a wavelength of 30 cm) signal fading and slow variation of large-scale fading have been observed both from the measurement and prediction results showing general agreement.

IV. CONCLUSION

In this paper, a novel frustum ray tracing technique has been introduced. Its ability of accommodating hundreds of thousands of receiving points in a complex building model has been demonstrated by a program code that runs on a typical PC. The prediction accuracy of narrowband channel parameters by the proposed technique has been verified by the measurement results. The proposed FRTT is shown to be useful in predicting channel characteris-

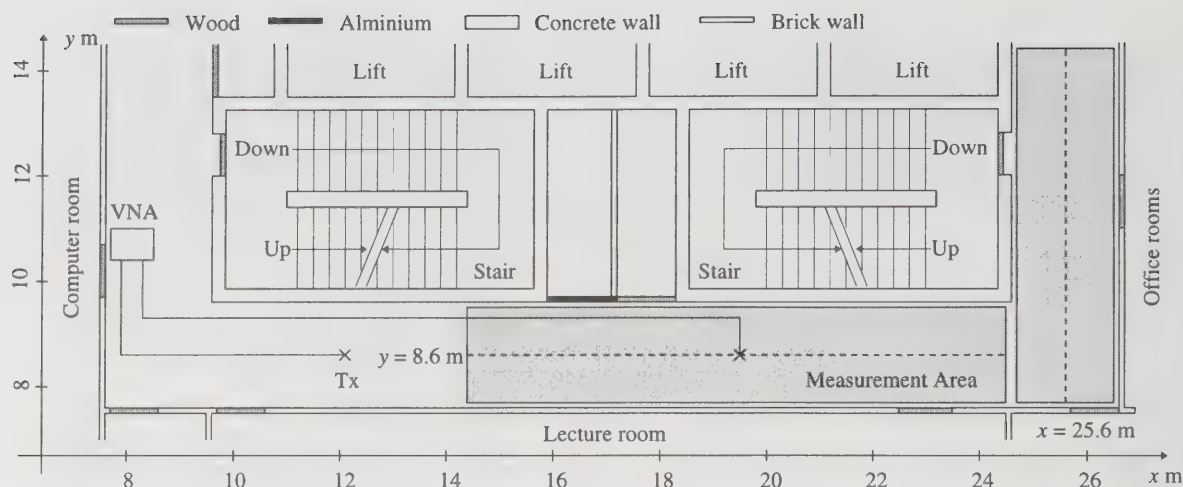


Figure 3: Measurement set-up.

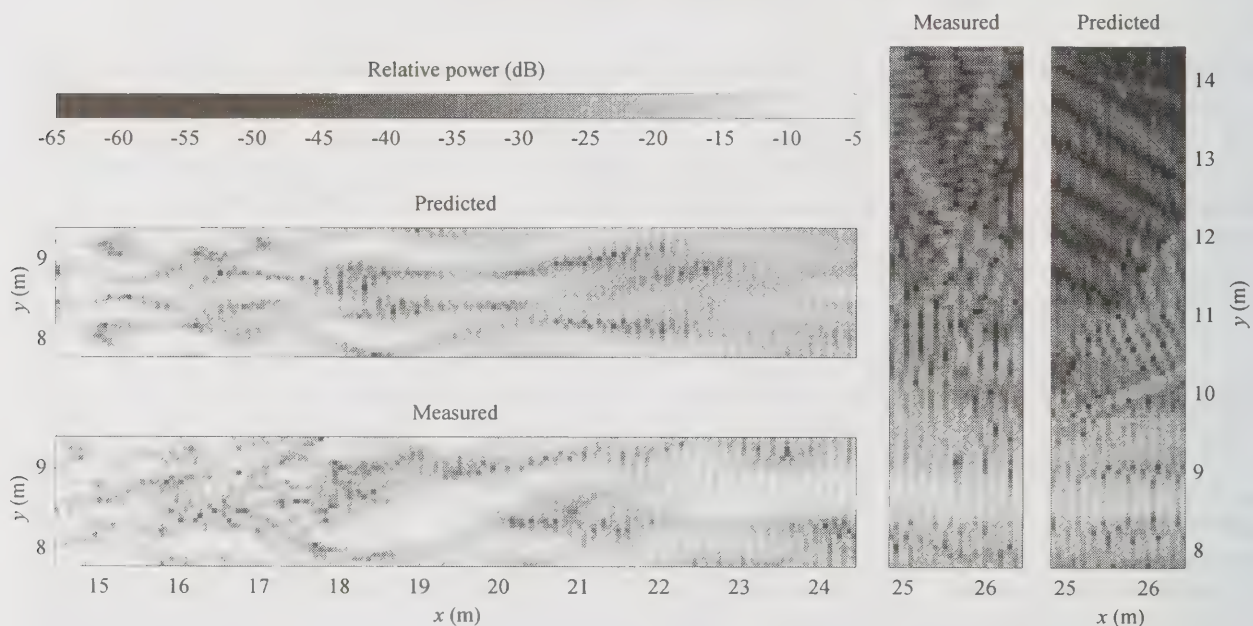


Figure 4: Measured and predicted received signal power map.

tic maps with high spatial resolution in any complex in-building environment.

REFERENCES

- [1] J. W. McKown and R. L. Hamilton, "Ray tracing as a design tool for radio networks," *IEEE Network Magazine*, vol. 5, no. 6, pp. 27–30, November 1991.
- [2] S. Y. Seidel and T. S. Rappaport, "Site-specific propagation prediction for wireless in-building personal communication system design," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 4, pp. 879–891, November 1994.
- [3] S-H. Chen and S-K. Jeng, "An SBR/image approach for radio wave propagation in indoor environments with metallic furniture," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 1, pp. 98–106, January 1997.
- [4] H. Suzuki and A. S. Mohan, "Ray tube tracing method for predicting indoor channel characteristics map," *Electronics Letters*, vol. 33, no. 17, pp. 1495–1496, August 1997.
- [5] D. F. Rogers, *Procedural Elements for Computer Graphics*, McGraw-Hill, 1985.

VHF Space-to-Earth Radio Link Performance in Various Man-Made Noise Environments

Robert J. Achatz and Roger Dalke
Institute for Telecommunication Sciences
Boulder, Colorado 80303
303-497-3498, 303-497-3680 (fax), rachatz@its.blrdoc.gov

Abstract

The performance of a differentially-coherent binary-phase-shift-keyed radio link operating at a frequency of 137 MHz in various man-made noise environments was determined by Monte-Carlo simulation. The man-made noise was generated by a model built from noise measurement data. The results show that the performance can be severely degraded in links that operate in residential and business environments or near power lines. Further measurements and analysis is needed to determine the extent of this degradation.

Key words: Man-made noise, space-to-earth radio link simulation, noise measurement and modeling.

Introduction

Satellite space-to-earth radio links in the 136- to 138-MHz VHF band, that use omnidirectional antennas at the earth terminal, are degraded by man-made noise unintentionally radiated from a variety of electrical and electronic devices. Services in this band susceptible to such degradation include mobile satellite paging and meteorological-satellite weather-image broadcasting. The man-made noise environment is characterized by the statistics of the average noise power (ANP) and instantaneous noise power (INP). Typically, statistics of man-made noise power used in today's satellite space-to-earth radio link design are derived from measurements taken more than 20 years ago [1,2]. Because electrical and electronic technology has changed over this period, the Institute for Telecommunication Sciences has performed "spot-check" measurements of the man-made noise to determine if statistics from the earlier noise measurements are still useful [3,4,5].

In this study, an INP model, built from the spot-check measurements, was integrated into the Signal Processing Workstation communication system simulation environment to predict the bit error ratio (BER) of a differentially-coherent binary-phase-shift-keyed (DCBPSK), space-to-earth radio link, using an omnidirectional earth terminal antenna, and operating in a man-made noise environment. The INP model produces noise whose amplitude probability distributions (APD) are similar to those encountered in our measurements.

Selected APD's

The APD's are plotted on a Rayleigh graph where the Rayleigh distributed amplitude of additive-white Gaussian noise (AWGN) forms a line with a slope of $-1/2$. The ordinate is in units of dB above kT_0b where k is Boltzman's constant, T_0 is 288 K, and b is the bandwidth, which was 30 kHz in our measurements. The abscissa is in units of percent of samples that exceed the ordinate. The selected APD's discussed in this section illustrate some of the distinctive characteristics of man-made noise in this band. One such characteristic is that the same location can have widely different INP statistics at different times. Figures 1a and 2a show APD's of the same residential location at two different times of the same day. The APD in Figure 1a has more frequent, lower amplitude, non-Gaussian, impulses than the APD in Figure 2a. Another characteristic is that the business environment measurements often exhibited both impulsive wideband and constant narrowband noise. The constant narrowband noise is thought to originate from periodic, pulsed emissions from nearby electrical or electronic equipment.

Figure 3a is an example of a business environment APD obtained in an office park setting. The slope of the APD at high probabilities is indicative of a Ricean-distributed noise amplitude due to constant narrowband noise. In this case, the ratio of constant-to-Gaussian noise power is 3 dB. Figure 4a shows an APD of noise measured near a power line in a rural environment. The APD shows that non-Gaussian impulses occur fairly frequently. Automobile noise impulses occurred much less frequently than power line noise impulses.

Simulation Results

The simulation consists of a DCBPSK radio link operating in various man-made noise environments at a 30 kbps data rate. Part 'b' of Figures 1-4 show the BER performance corresponding to radio link operation in the man-made noise environment whose APD is shown in part 'a' of these figures. The theoretical DCBPSK curve represents operation in AWGN. The error bars on the BER curves represent the standard deviation of the result due to a limited number of Monte Carlo trials.

The radio link margin is dependent on the signal-to-noise ratio (SNR) required for a specified BER. In this paper we assume a BER of 10^{-3} is required and use "margin degradation" to evaluate the performance of links operating in the different man-made noise environments. Margin degradation is the difference between the SNR in dB required to obtain a BER of 10^{-3} in the man-made noise environment and the 8 dB SNR required to obtain the same BER in AWGN. The approximate margin degradation for Figures 1b through 4b is 10, -5, 8, and 10 dB. At this BER there is no margin degradation for automotive noise.

Conclusions

Radio link margins for systems operating in AWGN can be predicted with the ANP alone. Our results show that radio link margins for systems operating in a non-Gaussian noise environment require additional statistics to determine performance, and that radio link margins determined with ANP alone are likely to be in error for many man-made noise environments at 137 MHz.

Previously, researchers investigating the BER in atmospheric non-Gaussian noise used the INP parameter V_d , defined as the ratio of the root-mean-square noise power to the ANP, to determine the required SNR and radio link margins [6-11]. Further work should determine the distribution of V_d in this band with respect to environment, time, and location. Our measurements also revealed the presence of constant narrowband noise in business environments. Further work is needed to determine the extent and source of this type of noise.

The measurements and models in this paper have only addressed first-order INP statistics. The power-line noise was found to exhibit periodic behavior which may affect the operation of radios using interleaving and error correction. Therefore it would be useful to measure and analyze the higher-order statistics of man-made noise in the future. Pulsed narrowband noise was

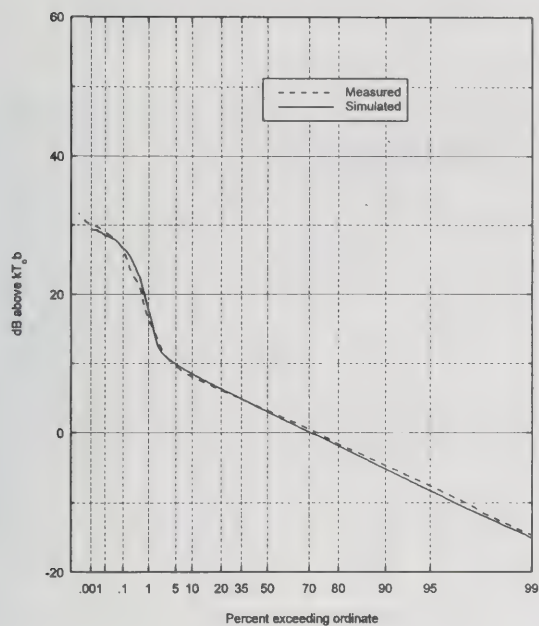
also observed and should be included in future measurement analyses.

Disclaimer

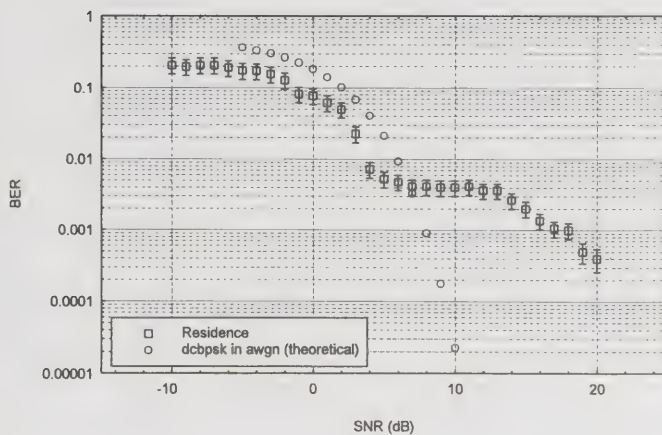
The Institute for Telecommunication Sciences does not endorse any commercial products mentioned in this paper.

References

- [1] CCIR, "Man-made radio noise," Report 258-5, International Telecommunications Union, Geneva, Switzerland, 1990.
- [2] W. Chrichlow, C. Roubiqu, A. Spaulding, and W. Beery, "Determination of the amplitude-probability distribution of atmospheric radio noise from statistical moments," J. Res. NBS, vol 64 D, Jan.-Feb. 1960, pp. 49-56.
- [3] R.J. Achatz, Y. Lo, P. Papazian, R. Dalke, and G. Hufford, "Man-made noise in the 136- to 138- MHz VHF band," NTIA Report (in publication).
- [4] R. Dalke, R. Achatz, Y. Lo, P. Papazian, and G. Hufford, "Measurements and analysis of man-made noise in VHF and UHF bands," in *Proc. of the 1997 Wireless Comm. Conf.*, Boulder, CO., Aug. 1997, pp. 229-233.
- [5] R. Dalke, R. Achatz, Y. Lo, "Statistics of man-made noise at 137 MHz," in *Proc. of the 1998 IEEE Conf. on Electromagnetic Compatibility*, Denver, CO., Aug. 1998, (in publication).
- [6] J.W. Modestino and B. Sankur, "Performance of selected linear and nonlinear receivers in ELF noise," in *Proc. of the 1974 ICC*, Minneapolis, MIN., 1974, pp. 30F-1 - 30F-5.
- [7] B. Shepelavey, "Non-Gaussian atmospheric noise in binary-data phase-coherent communication systems," *IEEE Trans. on Comm. Sys.*, Sep. 1963, pp. 280-284.
- [8] A.D. Spaulding, "Determination of error rates for narrow-band communication of binary coded messages in atmospheric radio noise," *Proc. of the IEEE*, Feb. 1964, pp. 220-221.
- [9] A.M. Conda, "The effect of atmospheric noise on the probability of error for an NCFSK system," *IEEE Trans. on Comm. Tech.*, Sep. 1965, pp. 280-283.
- [10] J.H. Halton and A.D. Spaulding, "Error Rates in Differentially Coherent Phase Systems in Non-Gaussian Noise," *IEEE Trans. on Comm. Tech.*, Oct. 1966, pp. 594-601.
- [11] A.D. Spaulding, "Digital system performance software utilizing noise measurement data," NTIA Report 82-95, Feb. 1982.

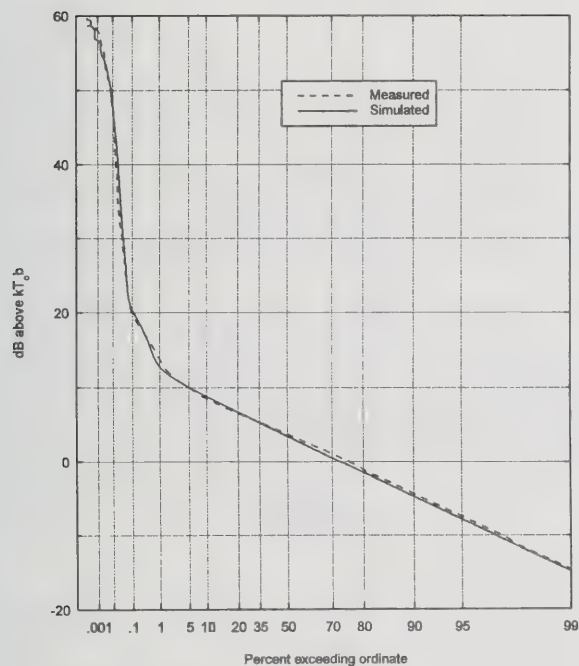


(a)

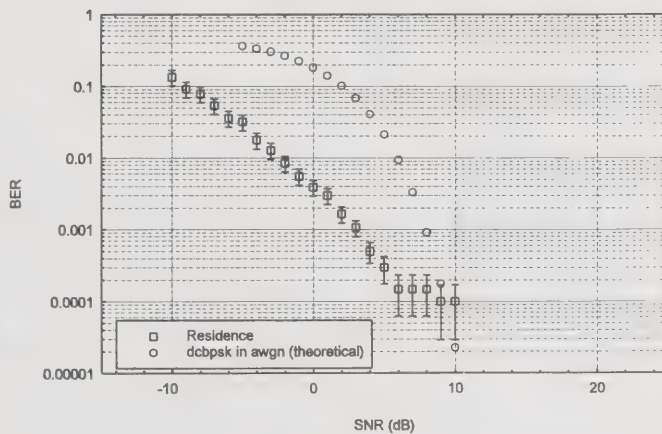


(b)

Figure 1. APD(a) and BER(b) of man-made noise at a Lakewood, CO, residence on November 10, 1996, from 12:00 a.m. to 12:40 a.m. Average noise power is 15.0 dB above kT_0b .

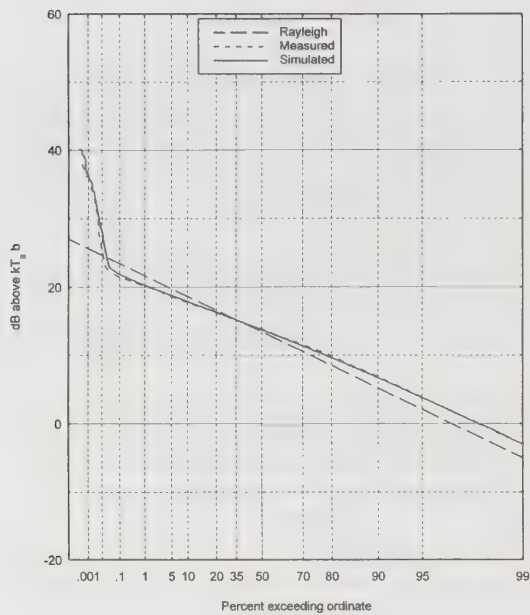


(a)

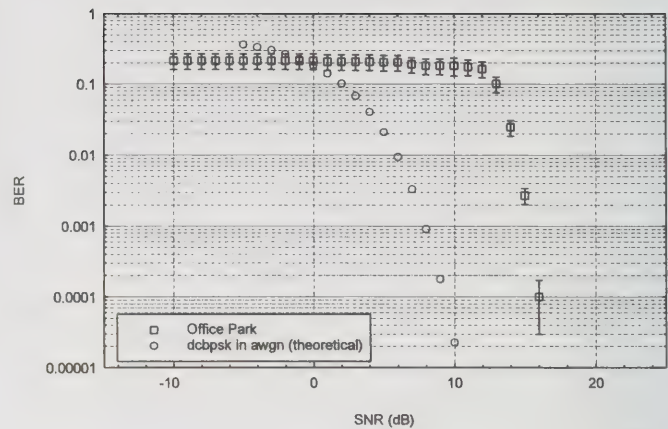


(b)

Figure 2. APD(a) and BER(b) of man-made noise at a Lakewood, CO, residence on November 10, 1996, from 3:30 p.m. to 4:00 p.m. Average noise power is 5.6 dB above kT_0b .

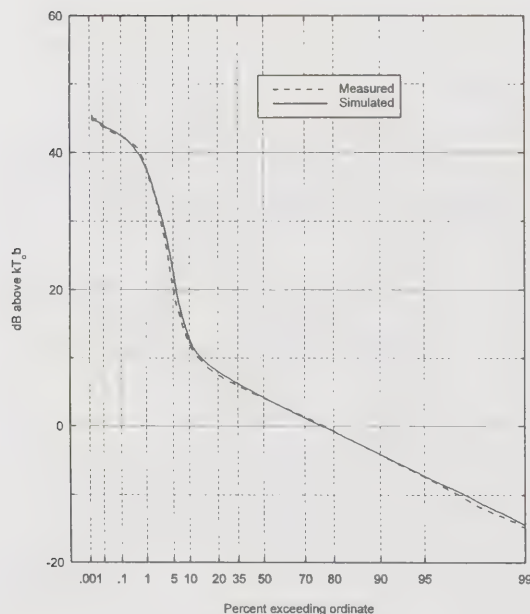


(a)

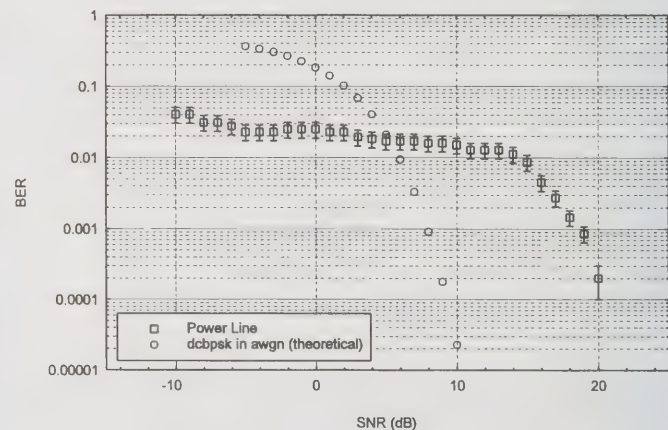


(b)

Figure 3. APD(a) and BER(b) of man-made noise measured in an office park in Golden, CO, on November 27, 1996, from 11:15 to 11:45 a.m. Average noise power is 14.5 dB above kT_0 .



(a)



(b)

Figure 4. APD(a) and BER(b) for power line noise measured in a rural environment 5 km west of Leyden, CO, on November 12, 1996, at 2:02 p.m. Average noise power is 22.6 dB above kT_0 .

Reception Algorithms for Ray Launching Modeling of Indoor Propagation

Santiago J. Flores, Luis F. Mayorgas, Francisco A. Jiménez
Departamento de Comunicaciones
Escuela Universitaria Gandia (Universidad Politécnica de Valencia)
Ctra. Nazaret-Oliva, s/n; 46730 Grao Gandia. (SPAIN)
Tel. +34-6-387 73 00. Fax +34-6-284 93 13
E-mail: sflores@com.upv.es

Abstract

Ray Launching is a deterministic method for modeling radio wave propagation based in the Geometric Optic theory. This method is very efficient in three-dimensional and complex environments like inside buildings. But some problems appear implementing the reception algorithms, because the receptors are considered punctual and the rays haven't thickness. The *Reception Sphere* method is a popular technique used to fix this problem. An alternative method is to launch three-dimensional tubes, instead of rays. In this paper, both techniques are analyzed, showing the advantages and disadvantages of each one in efficiency and accuracy.

Key words: Indoor Wireless Propagation Modeling, Ray Launching and Geometric Optics

Introduction

Ray Launching consists of the shooting of straight rays in the entire 4π stereoradians space, that propagate from the transmitter according to the Geometric Optics (GO) theory, and considering different additional propagation phenomena, like reflection, transmission, diffraction or scattering. The total received signal strength is calculated as the superposition of the magnitudes and phases of the individual rays arrived to the receiver. The principal advantage of this method is the very easy implementation in a computer [1].

Ray Launching Reception Algorithms

But with this modeling method, it's difficult to determine if a ray reaches a punctual receiver. A common method to resolve this situation is the *Reception Sphere* algorithm: If a ray intersects a reception sphere, with a defined radius (*the Impact Radius*) around a possible receptor, it is considered that it contributes to the total received signal.

For good result accuracy, it is desirable that rays maintain a constant angular separation from its neighbor. This ensures that each ray represents an equal, regularly shaped, and unique portion of the total wavefront. The impact radius of the sphere must effectively account the divergence of the rays from the transmitter. If is oversized, more than one ray from the same wavefront will be added wrongly to the total

received power. If it's undersized, the rays of the same wavefront will not reach some receivers (fig. 1).

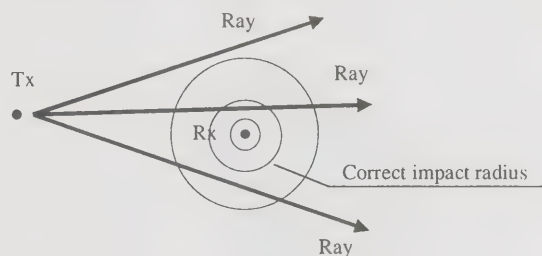


Figure 1. Reception sphere and impact radius

In two dimensions, this problem can be easily solved. For two neighboring rays separated an angular spacing α at the transmitter, and at a distance d from the transmitter, the linear distance between them is approximately $L=\alpha d$. So, we can take $r = \alpha d/2$ as the impact radius (fig. 2).

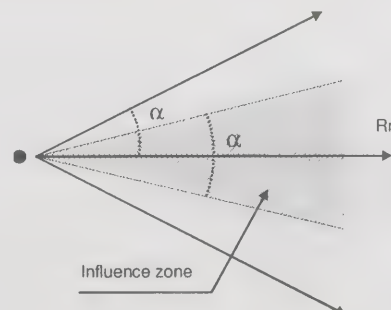


Figure 2. 2D influence zones

But in 3D situations, the ray influence zone is a three-dimensional cone, whose vertex is the transmitter point and whose axis is along the ray path (fig. 3).

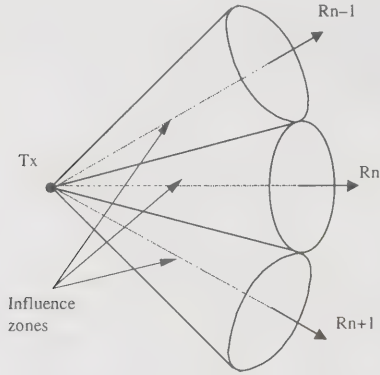


Figure 3. 3D influence zones

It seems intuitive to take $r = \alpha d/2$ again, but some problems will arise. Let's consider a cross section at a distance d from the ray origin, perpendicular to the Rn ray. Reception spheres cross one to others and it appears some *shady zones* that don't belong to any influence zone. No rays will arrive at any receiver contained in these shady zones (fig. 4).

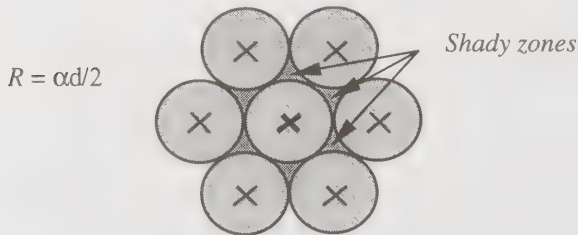


Figure 4. Shady zones

This problem can be solved increasing the radius to cover all the shady zones. Taking $r = \alpha d/\sqrt{3}$, all of them can be covered. However, some influence zones will overlap other, appearing multiple rays from the same wavefront (fig. 5).

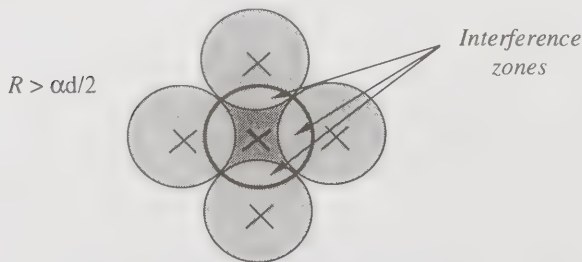


Figure 5. Interference zones

This trouble is called *double count error* and produces a maximum of +6dB power error per ray. Considering an ideal wavefront with the rays distributed uniformly with a hexagonal geometry, the probability of a double counting occurs is:

$$\frac{2\pi}{3\sqrt{3}} - 1 \approx 21\%$$

So, the mean received power will be increased by at least 1.26dB, with an additional standard deviation of ± 2.4 dB. Actual simulations show worst results (fig. 6).

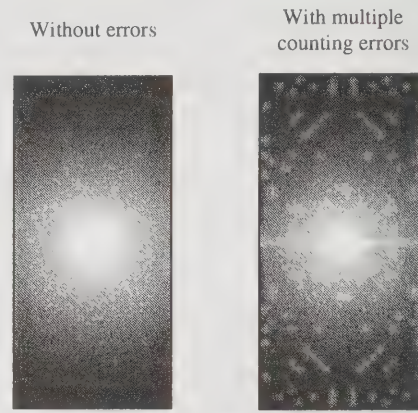


Figure 6. Effect of the multiple counting errors in the received power

Double count errors could be fixed with an algorithm that, for each received ray, verifies if there's any other previous received ray with similar direction and the same number of transitions (reflections or transmissions). If this happens, the new received ray may be redundant with the previous, and it hasn't to be considered. This method is called *brute-force double ray counting verification*.

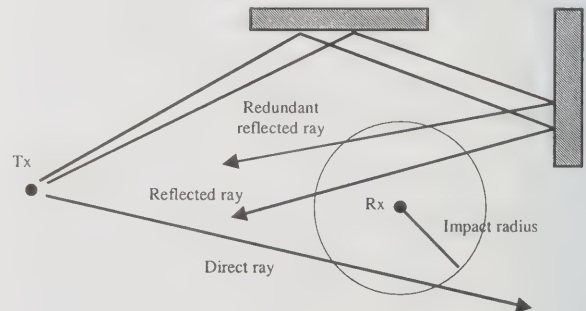


Figure 7. Two rays with the same number of reflections and transmissions and similar direction are redundant

Another possibility is to use a weighting function that weights the power level of each received ray, so that the redundant contributions will be compensated. This method is called *distributed wavefront technique* and remedies well the double counting errors in open environments [2], but it shows some troubles in 3D indoor scenarios, and it spoils the power delay profiles in wideband simulations.

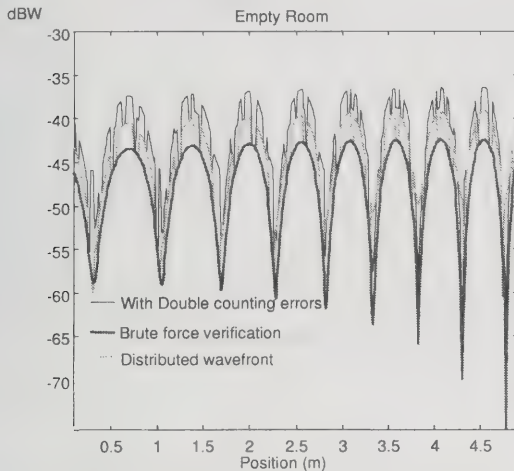


Figure 8. Performance of double counting elimination methods in an empty room

So, figures 8 and 9 show the graphics that represent two indoor simulations. Both methods of double counting errors fixing are compared with the results obtained without fixing. The above simulation was made in an empty room, considering only one reflection on the walls, the floor and the ceiling. The distributed wavefront method reduces some double counting errors, but not all of them. Mean received power is reduced in 2.07dB with this method, in front of the 5.1dB reduced by brute force.

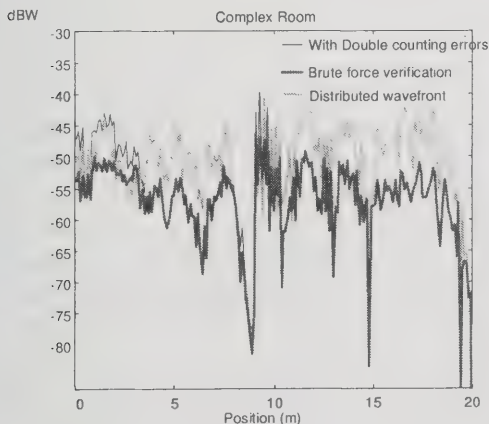


Figure 9. Performance of double counting elimination methods in a furnished room

The simulation of the fig. 9 was made in a more complex indoor environment with some furnitures and plasterboards, and considering not only reflections, but also transmissions and diffraction. In that situation, distributed wavefront method can only reduce the mean received power only in 0.72dB, in front of the 6.29dB reduced by brute force. CPU times are not critic in the election of the method, because in both cases are very similar.

Tube-Launching Reception Algorithms

Tube Launching is a variation of Ray Launching method, where the structures launched from the transmitter have three dimensions: tubes. One tube is made with several rays that limit the tube (fig. 10).

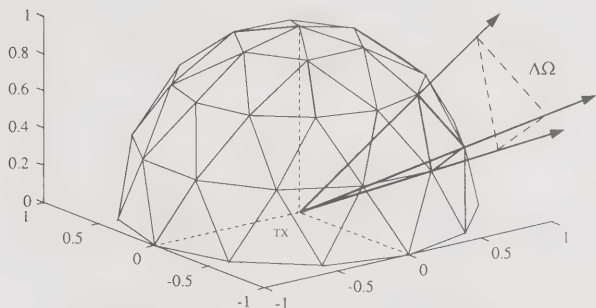


Figure 10. Tube generation and launching

With this method it's unnecessary the impact radius algorithm, because the power associated to the tube is added to the receptors located inside it. Influence zones for adjacent tubes will not intersect between them, doing unnecessary to check the received impacts to verify if they are redundant [3].

This is the main advantage for tubes, especially with high number of transitions (reflections or transmissions). The processing of received impacts consumes a great mount of computer time. Indeed, profuseness of reflections or transmissions results on more impacts on the receivers. There's no extra time increment in simulation time for Tube Launching. However, in Ray Launching, the received ray list will be larger for each receiver, and more comparison time will be required for each new received ray, increasing the simulation time.

On the other hand, behavior of tubes is different than the rays when an obstacle is reached. So, let's consider a metallic surface between the Tx and the Rx, near to the Rx. The wavefront will not directly arrive to the Rx. There's no problem for Tube Launching. Tubes are lanced from the Tx, until they

intersect the metallic surface. Its sides and the metallic surface (fig. 11a) limit the influence zone.

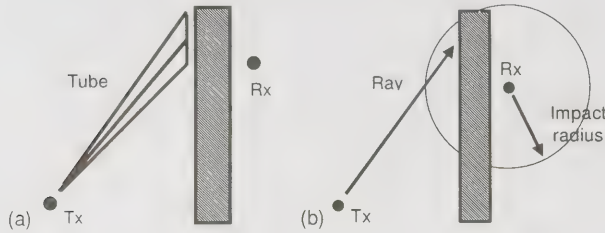


Figure 11. Tubes (a) and Rays (b) in front of surfaces

However, for Ray Launching, the rays from the T_x will intersect the metallic surface too, but the algorithm compares distance between the ray and the R_x , and the impact radius, and then considers that the ray hits the R_x (fig. 11b). This undesired effect is very important when the impact radius is large (low resolutions or large distance from T_x). In fig. 12, the distortion around metallic surfaces for rays can be observed, in comparison with the almost perfect definition for tubes. Both simulations were made in a rectangular room, with two metallic walls crossing in the middle.

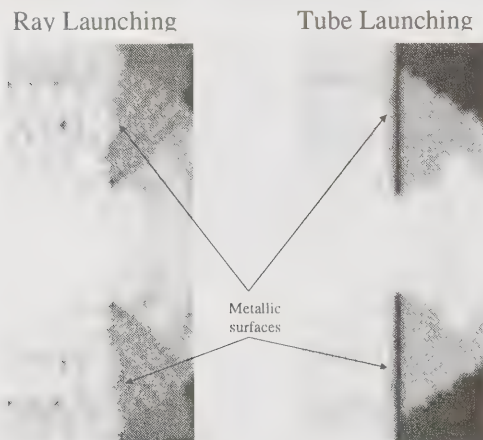


Figure 12. Room with two metallic surfaces

A similar problem arises on places next to the corners. The effect of the reflected wavefront on corners must be considered. With Tube Launching, those tubes that meet any corner along its path will have some axis that will intersect with one surface and other axis will intersect on the other surface. If the tube splits into two (one for each surface), it will result on more execution time. It's better to reject the reflected tube. But then, a shady zone will appear near the corners, in those points not reached by the reflected tubes (fig. 13).

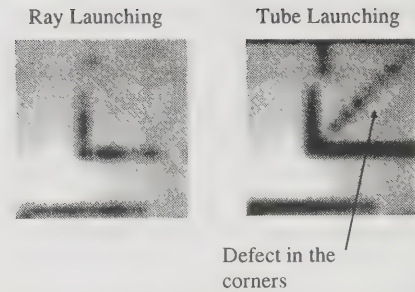


Figure 13. Simulation in a room with corners

Tubes are also problematic in complex environments with high dimensions, like office buildings. In those cases, tubes will travel large distances and it's probably that an obstacle will intercept only a little fraction of the tube, involving that the tube has to be divided in several, spending more simulation time. Alternatively, the tube could be discarded, resulting on worst results.

Conclusions

Ray Launching method has a serious problem with double counting errors. In indoor environments, these errors can only be perfectly fixed using *brute force* verification.

Another solution is to use Tube Launching, that decreases the simulation time, especially for environments with a high level of transitions. Tubes are also better in places close to thin surfaces like partition walls. However, Ray Launching could work better in complex environments with high dimensions and with a great mount of obstacles, surfaces, and corners. A sensible choice will depend on the specific simulation environment.

References

- [1] D.J. Cichon, W. Wiesbeck; "Indoor and outdoor propagation modeling in pico cells", Proceedings of PIMRC'94, pp.491-495. The Hague, Netherlands, September, 1994
- [2] Greg Durgin, Neal Patwari, Theodore S. Rappaport; "An Advanced 3D Ray Launching Method for Wireless Propagation Prediction", Proceedings of VTC'97. Phoenix, USA, may, 1997
- [3] S.J. Flores, L.F. Mayorgas, F.A. Jiménez; "Two Deterministic Techniques for Indoor Propagation Modelling: Tube Launching and Ray Launching", Proceedings of ICT'98. Porto Carras, Greece, jun, 1998

Design Considerations for Extremely High-Q Integrated Inductors and Their Application in CMOS RF Power Amplifier

Tony Yeung, Jack Lau, H. C. Ho, M. C. Poon

Dept. of EEE, The Hong Kong University of Science & Technology

Clear Water Bay, Hong Kong

852-235-88843, 852-235-81485 (fax), eetony@ee.ust.hk, eejlau@ee.ust.hk

Abstract

An extremely high Q monolithic inductor ($Q > 2000$) on silicon substrate was reported recently [1]. The reported Q is 3 order of magnitude higher than previously reported monolithic inductor [2]. Such high quality factor may greatly improve the performance of monolithic RF circuits. In this paper, both 1-Port (one terminal at ground) and 2-Port (no terminal at ground) scattering parameters of the high- Q inductor were compared to examine any possible differences in device characteristics. A broadband physical model of the active inductor on silicon are presented to illustrate the improvement of quality factor. A design space with a range of gain and phase difference of current in the coils of the active inductor is reported. A typical application of the active inductor in CMOS RF power amplifier design is shown to illustrate the feasibility of applying the active inductor in RF circuits design.

I. INTRODUCTION

Nowadays, as the demand for wireless communication is increasing tremendously, high quality on-chip inductors are very important for high level integration of RF circuits in order to reduce costs. Typical applications of on-chip inductors include low loss inductors for input matching of low noise amplifier [3], inductively loaded pre-amplifiers and output match for high efficiency switch-mode power amplifiers [4], and high- Q tank circuits for low phase noise voltage control oscillators [5]. Recently, many research activities have been focused in the design, model and optimization of the spiral inductor on silicon substrate [6][7]. However, the reported Q are limited to below 100 at GHz range, and recently, an extremely high- Q monolithic inductor ($Q > 2000$) on silicon substrate was reported [1]. Such a high quality factor may greatly improve the performance of monolithic RF circuits. Yet, while the report focused mostly in 1-Port experimental verification, the optimal parameters of the supporting circuits and the feasibilities in RF application have not been reported. In this paper, both 1-Port (one terminal at ground) and 2-Port (no terminal at ground) scattering parameters of the active inductor were compared to examine any possible differences in device characteristics. A suitable lumped broadband equivalent circuit model of the active inductor are presented to illustrate the improvements in quality factor and can be incorporated into circuit simulators e.g. SPICE for time domain simulations. A design space with a range of gain

and phase difference of the current in the coils of the active inductor is reported. A typical application of the active inductor in CMOS RF power amplifier design is shown to illustrate the feasibility of applying the active inductor in RF circuits design. A simulated drain efficiency of 80% is obtained for a 835MHz Class E power amplifier in 0.8 μ m CMOS technology.

II. LUMPED EQUIVALENT CIRCUIT MODEL

The inductor, which is essentially a 2-Port transformer, consists of a primary coil (IN1, OUT1) and a secondary coil (IN2, OUT2) is shown in Fig.1. The objective here is to use the secondary coil as a conduit to compensate for the energy loss in the primary coil. The self-inductance of the primary coil can be varied by changing

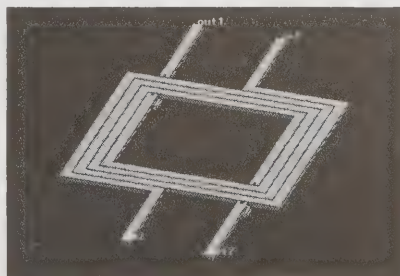


Fig. 1 A 2-Port transformer forms the inductor. IN1, OUT1 forms the primary coil and IN2, OUT2 forms the secondary coil. The size of the transformer is 500 μ m by 500 μ m with 2 turns of 15 μ m of metal width. High Q is achieved by compensating energy loss of the primary coil with energy injection through the secondary coil.

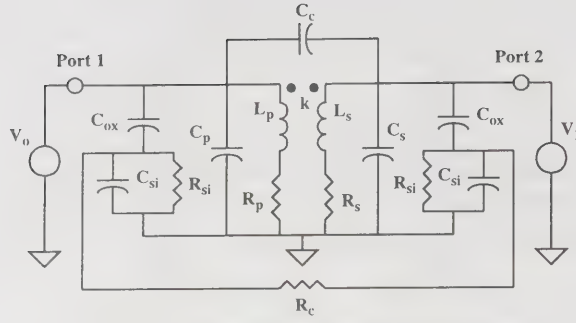


Fig. 2 Physical lump model of the transformer on silicon substrate.

the mutual coupling between the primary and secondary coil. Fig. 2 shows the broadband equivalent circuit lump model of the transformer on silicon substrate. This model is the extension of the physical model for planar spiral inductor on silicon [6]. C_{ox} , C_{si} and R_{si} model the substrate injection at high frequency. L_s , R_s and C_s represents the series inductance, resistance and feed-capacitance respectively. C_c models the capacitive coupling through the oxide between the two ports and R_c models the substrate coupling between the inductors. The mutual coupling factor between the two inductors is model by k .

By observing the input impedance of Port 1 of the transformer and varies the phase of voltage applied on Port 2 i.e. $V_1 = V_0 \angle \theta$, one can observe that the highest mutual inductance can be obtained when current of same frequency and phase are injected into Port 1 and Port 2. While, the lowest mutual inductance is obtained when out of phase (180°) current flow into two ports. By varying the relative phase of current between primary and secondary coil, the self-inductance of the primary inductor will vary, but the quality factor won't be improved by changing the relative phase alone. A variable gain amplifier can be used to amplifier the current signal before feeding to the secondary coil, i.e. $V_1 = AV_0 \angle \theta$. Fig. 3 shows the setup to determine the 1-Port and 2-Port char-

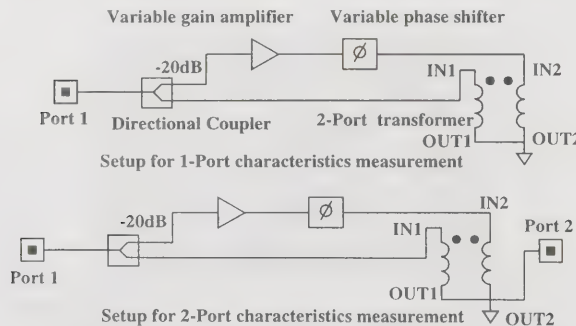


Fig. 3 Setup for 1-Port and 2-Port characteristics measurements of the inductor. Port 1 is connected to the primary coil through a directional coupler and to the secondary coil through a directional coupler, a variable gain amplifier and a variable phase shifter.

acteristics of the active inductor. The primary coil is connected to Port 1 through a directional coupler, and the secondary is connected to Port 1 through the directional coupler, a variable amplifier and a variable phase shifter. The difference between 1-Port and 2-Port characteristics measurements is that OUT2 is connected to Port 2 for 2-Port measurements.

III. SIMULATION METHODS

An electromagnetic simulation tool, Sonnet [8] was used to extract scattering parameters of the 2-Port transformer. During simulations, we have included the effects such as substrate loss due to the high frequency coupling to the low resistivity substrate. After acquiring the s-parameters of the 2-Port transformer, performance simulation of the active inductor was done by including the amplifier and phase shifter in circuit simulation tools with Libra in EEsof [9]. Fig. 4 shows the 1-Port S_{11} parameter of the primary inductor at 1GHz with constant amplifier gain of 22.6dB and the relative phase of power between the inductors varies from 0° to 360° . The high-

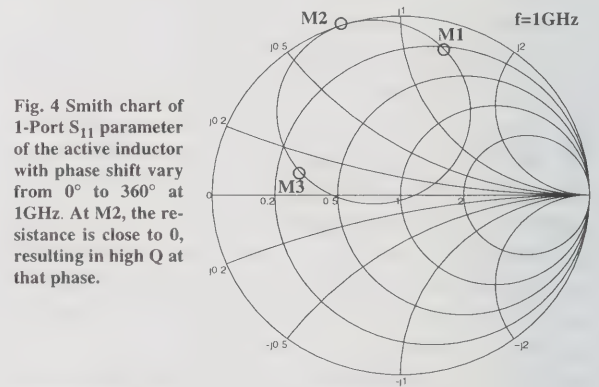


Fig. 4 Smith chart of 1-Port S_{11} parameter of the active inductor with phase shift vary from 0° to 360° at 1GHz. At M2, the resistance is close to 0, resulting in high Q at that phase.

Marker	Phase	Resistance	Reactance	Q
M1	0°	12.020 Ω	65.765 Ω	5.471
M2	25°	0.024 Ω	36.557 Ω	1523
M3	180°	11.887 Ω	8.188 Ω	0.689

est Q (=1523) is achieved with 25° phase difference in power (Marker M2). Moreover, with a phase difference of 0° (Marker M1), the inductance of the primary inductor is more than double of its self-inductance. A small inductance is found when the phase difference is 180° (Marker M3). The reactance can also be capacitive at some phase difference. Fig. 5 shows how the inductance and resistance changes as the phase difference varies with the amplifier gain of 22.6dB. The size of the transformer is $500\mu\text{m}$ by $500\mu\text{m}$ and the self inductance of the primary inductor is 4nH. As shown in the figure, the inductance can vary from 0 to 13nH and resistance is close to 0 at 25° .

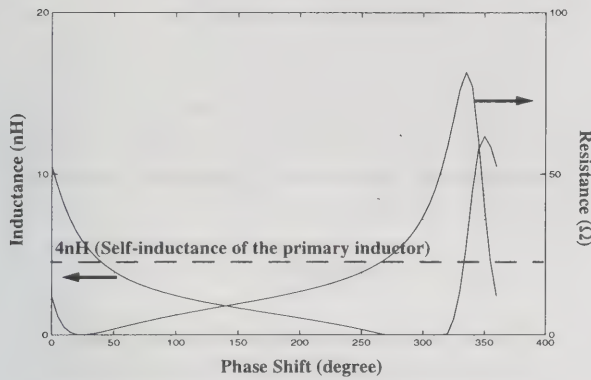


Fig. 5 Inductance and resistance of the primary inductor. Inductance can vary from 0 to 13 nH and resistance is close to 0 at 25°.

In order to give a better visualization of how the quality factor and inductance change with gain of the amplifier and phase of the phase shifter, we show the 3D plotting of Q and inductance with different phase and gain in Fig. 6 and Fig. 7. As observed from the figures, at 1 GHz, high Q of more than 1000 is confined to a small combination of gain and phase. Also, the inductance is low for phase around 180°. From these two figures, we can obtain the requirement of phase and gain for designing the active inductor to meet the specified requirements on Q and inductance value. Fig. 8 shows the 1-Port and 2-Port S_{11} parameters from 0.1 to 6 GHz of a active inductor as compared with a standard spiral inductor. The standard spiral inductor is simply the primary coil of the transformer on silicon without the secondary coil. The figure shows that higher Q can be obtained by the active inductor for both 1-Port and 2-Port configuration in a wide range of frequencies with a constant amplifier gain of 22 dB and phase shift of 25°. Moreover, the inductor behaves differently in 2-Port configuration such that the inductance is smaller than that achieved in 1-Port configuration for the active inductor.

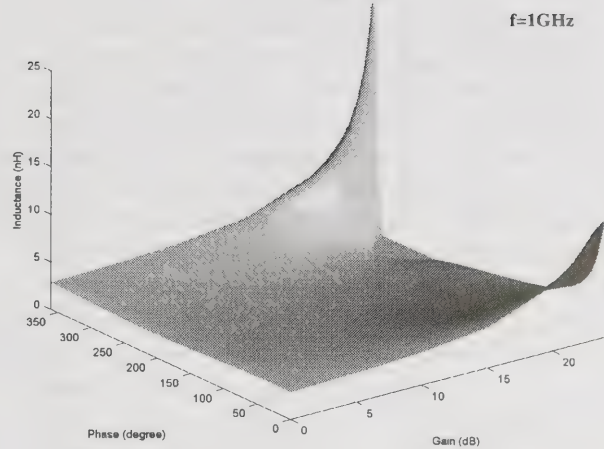


Fig. 7 Inductance plot as a function of gain and phase difference between power of the 2 inductors. The tunability of the inductance is large, and large inductance can be obtained at high gain.

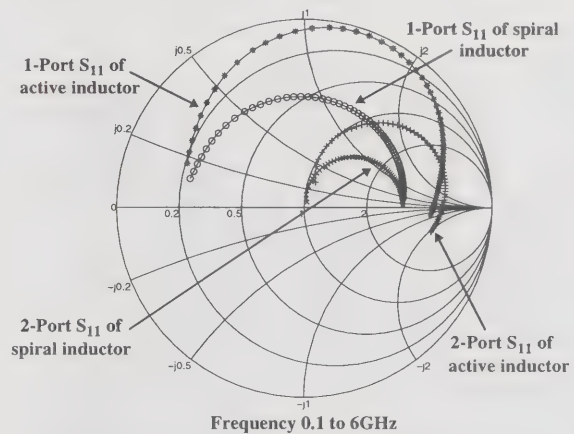


Fig. 8 Smith chart of 1-Port and 2-Port S_{11} parameters of a active inductor and a standard spiral inductor. Higher Q can be obtained by the active inductor for a wide range of frequency.

IV. APPLICATION OF THE ACTIVE INDUCTOR

In RF power amplifier, a high Q inductor helps improving power efficiency. The simulations as mentioned above can help us design the active inductor for use in RF power amplifiers. A Class-E power amplifier at 835 MHz has been designed with the active inductor as part of the circuit. Fig. 9 shows the overall schematic for the whole power amplifier, which includes a harmonic amplifier as a driver stage to generate square-like waveform for the power transistor, a Class-E switch mode power amplifier, an active inductor (with variable gain amplifier and phase shifter) and an output matching network in transforming the standard 50 Ω load of the antenna to the desired optimum load R_L . By varying the phase and gain the of the active inductor to the optimum value, high Q can be obtained in the primary inductor to increase the drain efficiency of the Class-E power ampli-

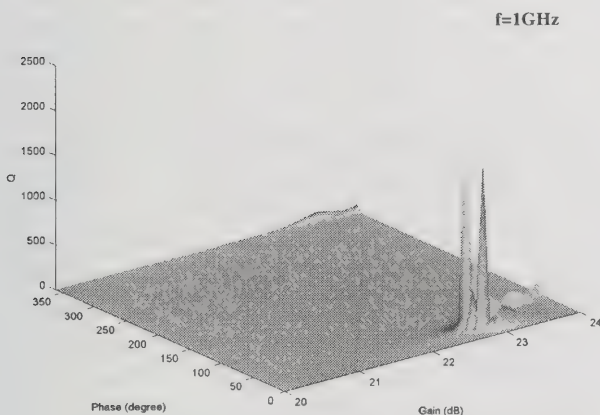


Fig. 6 Quality factor of the primary inductor. High Q can be obtained by adjusting the phase and gain, but it is confined to a small combination of gain and phase.

er and the output matching network is a simple T-network with part of the inductance being absorbed to the 2-port transformer. Table 1 summarizes the operation parameters of the Class-E switch-mode power amplifier, and by using the broadband lump model of the transformer obtained in previous section. SPICE transient analysis of the power amplifier can be achieved. The simulated drain efficiency is about 80%. A prototype of the CMOS power amplifier is under fabrication to verify the feasibility of using the active inductor in high frequency power amplifier design.

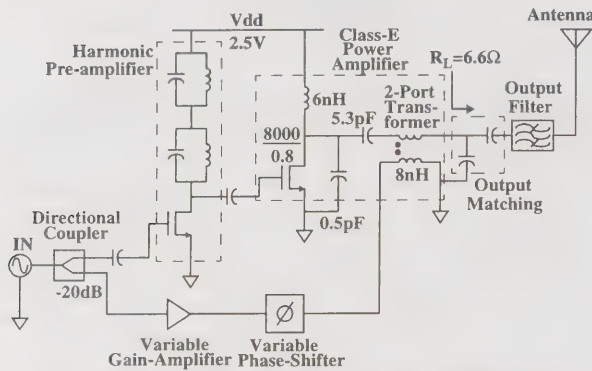


Fig. 9 Application of the active inductor in power amplifier design. The power amplifier include a harmonic pre-amplifier in the input stage for square wave output, a Class-E switch-mode power amplifier, input (not shown) and output matching network and the active inductor (with variable gain amplifier and phase shifter).

Table 1: Parameters of Class-E power amplifier

Frequency	835MHz
Technology	CMOS 0.8 μ m
Output Power	0.5W (27dBm)
Supply Voltage	2.5V
Simulated Drain Efficiency	80%

V. CONCLUSION

A broadband physical model of the active inductor on silicon are presented to illustrate the improvement of the quality factor by the active inductor. Both 1-Port and 2-Port scattering parameters of the active inductor were compared to examine any possible differences in device characteristics. A design space with a range of gain and phase difference of the current in the coils of the active inductor is reported. A typical application of the active inductor in CMOS RF power amplifier design is shown to illustrate the feasibility of applying the active inductor in RF circuits design.

VI. ACKNOWLEDGMENTS

The authors would like to thank the UMAN group for their assistance, especially to Alan Pun, Louis Tsui, Frankie Hui, Zhiheng Chen and Liu Wing Faat for their invaluable help and discussion. This project was supported in part by the RGC grant.

REFERENCES

- [1] D. R. Pehlke, A. Burstein and M. F. Chang, "Extremely High-Q Tunable Inductor for Si-Based RF Integrated Circuit Applications," *Proceedings of the IEEE International Electron Device Meeting*, pp. 63-66, 1997.
- [2] J. N. Burghartz, A. E. Ruehli, K. A. Jenkins, M. Soyuer and D. Nguyen-Ngoc, "Novel Substrate Contact Structure for High-Q Silicon-Integrated Spiral Inductors," *Proceedings of the IEEE International Electron Device Meeting*, pp. 55-58, 1997.
- [3] D. K. Shaffer and T. H. Lee, "A 1.5V, 1.5GHz CMOS Low Noise Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 5, pp. 745-759, May, 1997.
- [4] T. Sowlati, A. T. Salama, J. Sitch, G. Rabjohn and D. Smith, "Low Voltage, High Efficiency GaAs Class E Power Amplifiers for Wireless Transmitters," *IEEE Journal of Solid-State Circuits*, vol. 30, no.10, pp. 1074-1080, Oct., 1995.
- [5] J. Craninckx and M. Steyaert, "A 1.8GHz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 5, pp. 736-744, May, 1997.
- [6] C. P. Yue, C. Ryu, J. Lau, T. H. Lee and S. S. Wong, "A Physical Model for Planar Spiral Inductors on Silicon," *Proceedings of the IEEE International Electron Device Meeting*, pp. 155-158, 1996.
- [7] A. M. Niknejad and R. G. Meyer, "Analysis and Optimization of Monolithic Inductors and Transformers for RF ICs," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 375-378, 1997.
- [8] "Sonnet User's Manual," *Sonnet Software*, Sep., 1996.
- [9] HP-EEsof Microwave & RF Design, Series IV, 1996.

Microfabricated Single Crystal Silicon Transmission Lines

T.D. Kudrle, H.P. Neves, N.C. MacDonald

Phillips Hall

Cornell University

Ithaca, NY 14850

607-255-7377, 607-255-8601 (fax), tdk@ee.cornell.edu

Abstract

A novel approach to the transmission of microwave energy is demonstrated on standard (resistivity=1-20 ohm-cm) silicon. The transmission lines consist of pairs of parallel-plate waveguides, each waveguide formed from two deep (150 μm) suspended single crystal silicon (SCS) beams. This transmission is achieved through the integration of SCREAM (Single Crystal Reactive Etching and Metallization) technology adapted for very high aspect ratio structures with thick copper sputter deposition. The integration of these processes allows for the deposition of thick ($>.5 \mu\text{m}$) metal on the sidewalls and attenuation characteristics better than 0.17 dB/mm over the 10-48 GHz frequency range.

Introduction

Recent advances in microwave technology have yielded high performance transmission lines that are compatible with other microelectronics technologies [1]. There have also been advances in the fabrication of high-performance, high aspect-ratio micromachined transmission lines [2].

Transmission lines with tunable impedances have also been introduced. This tuning was achieved through the integration of the transmission lines with electrostatic actuators made possible with the SCREAM process. The spacing of these particular lines determines their impedances; therefore, by moving the lines with respect to one another, the line impedances can be controlled. Impedance tuning of 20 ohms about the 50 ohm center impedance was demonstrated with a bias of 70 V [3]. The lines offer exciting potential for a new class of microwave passive devices such as variable attenuators, tunable filters, etc.

The lines do however suffer from two major drawbacks: relatively poor attenuation characteristics ($>.6 \text{ dB/mm}$ at 40 GHz) and the reliance on high-resistivity ($>2000 \text{ ohm-cm}$) silicon. This paper reports on efforts to eliminate these two limitations. The improvements are achieved mainly through the use of fluorine-based etching technology. This technology permits the fabrication of much deeper structures that yield smaller ohmic losses in the transmission lines. The integration of this deep-etching technology with the thick deposition of Cu has yielded much improved results.

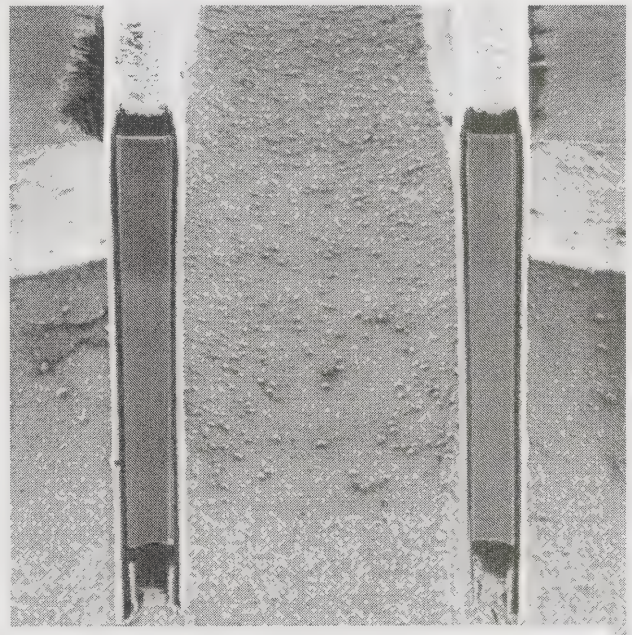


Figure 1. SEM micrograph of cross-section of parallel plate waveguide.

Design of Transmission Lines

The transmission lines can be viewed as consisting of two sets of parallel-plate waveguides in parallel with one another. A cross section of a fabricated set of 50 μm deep lines taken with a focused ion beam system is shown in Figure 1. The characteristic impedance of a pair of lines is thus simply half of the impedance of each individual waveguide. The characteristic impedance of a parallel-plate

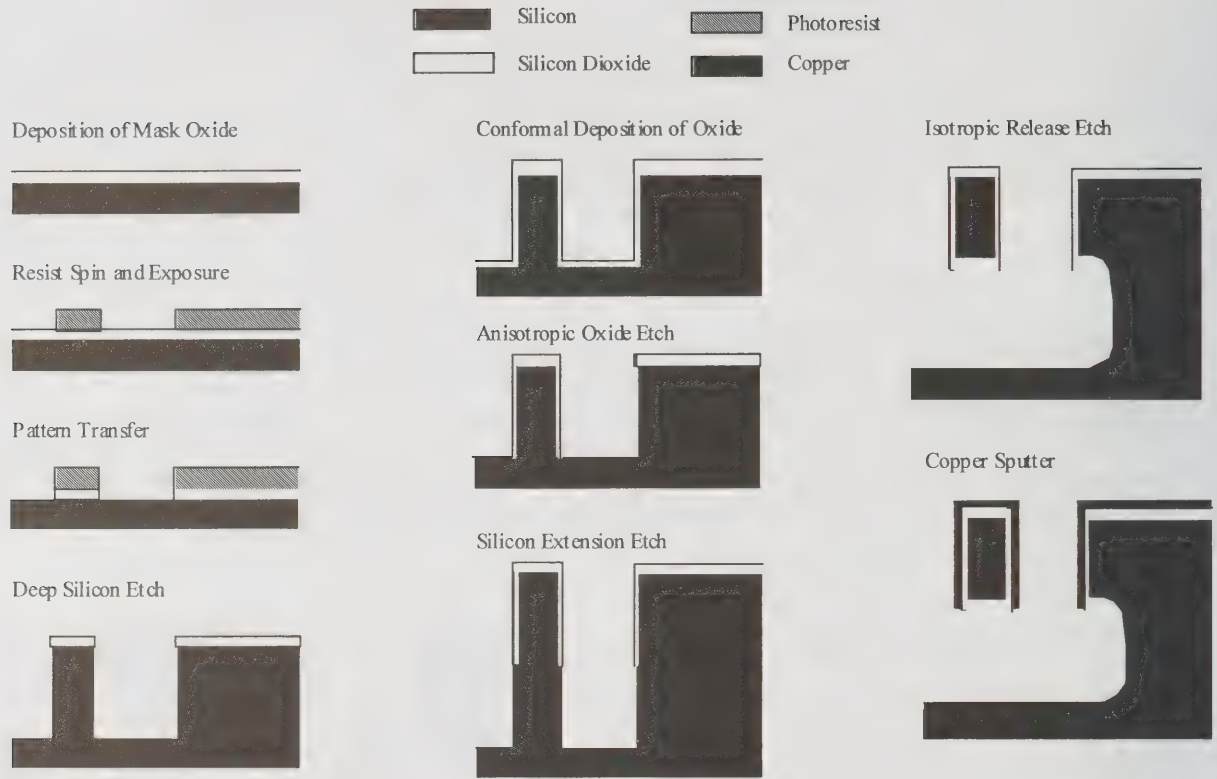


Figure 2. Diagram of major steps in SCREAM process flow.

waveguide can be determined from its approximate capacitance [4],

$$C = \epsilon \frac{h}{s} \left[1 + \frac{s}{\pi h} \left(1 + \ln \frac{2\pi h}{s} \right) \right]$$

where h is the height of the plates and s is the spacing between them. The impedance can then be determined from

$$Z_0 = \frac{\sqrt{\mu\epsilon}}{C}.$$

For transmission lines 50 μm deep, a 50 ohm impedance is achieved with a line spacing of 20 μm . This spacing can be varied by the integration of the lines with electrostatic actuators as has been demonstrated in [3] for lines 10 μm deep.

Fabrication of Transmission Lines

The transmission lines consist of suspended SCS beams coated with metal. These beams are fabricated

using a modification of the SCREAM process developed at Cornell University. SCREAM is a single-mask, low temperature process that can be implemented on fully processed VLSI wafers [5]. Devices fabricated with the SCREAM process include accelerometers, loading devices, and scanning tunneling microscopes. The process flow is shown in Figure 2 showing a fully released beam and an anchor. After depositing a layer of silicon dioxide, the wafer is spun with photoresist and the pattern is transferred to the mask oxide. The wafers are then etched in a Plasma Therm 770 fluorine-based reactive ion etcher (RIE) to a depth of 50 μm . A second layer of silicon dioxide is then deposited conformally on the structures and anisotropically etched back from the mesa tops and floors leaving a layer of protective oxide on the mesa tops (the remainder of the original mask oxide) and on the sidewalls. The structures are then placed in the RIE where they are etched again and then released from the substrate. The vertical distance between the bottoms of the beams and the floor is determined by length of the pre-release etch coupled with the length of the release etch itself. This floor lowering is important because it allows greater impedance control of the lines.

The wafers are coated with a thick sidewall oxide prior to metallization to reduce capacitive coupling to the Si beam cores. A number of different metallization techniques were

investigated including electroless Cu deposition, Cu sputtering, and Al sputtering. The sputtering of Cu yielded lines with the best characteristics and the data for these lines are presented in this paper. The sputtering was performed with a DC magnetron sputter system. Figure 3 shows an SEM micrograph of a fully fabricated transmission line with launch pads.

Measurements

All microwave measurements were performed with the HP 8510B network analyzer. The transmission lines were calibrated with the TRL technique using thru, reflect, and line standards [6]. These standards were fabricated along with the devices under test on each die. Before each set of measurements, calibration was performed on the structures to allow for the de-embedding of DUT data.

Results of attenuation for 150 μm deep lines fabricated on standard-resistivity silicon are shown in Figure 2. The attenuation is less than 0.18 dB/mm over the frequency range from 10 to 50 GHz.

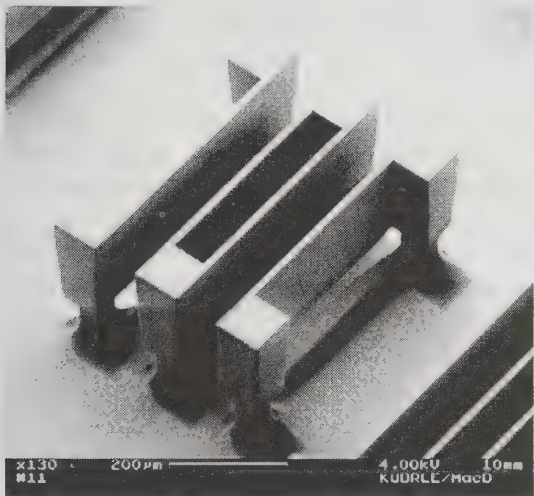


Figure 3. SEM micrograph of thru transmission line for TRL calibration.

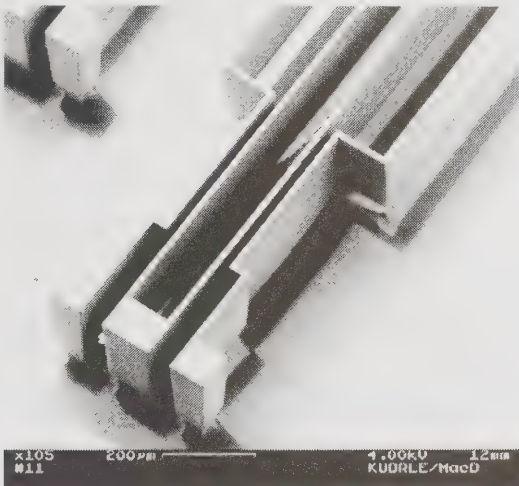


Figure 5. SEM micrograph of first three sections of stepped-impedance filters. The impedances, from the launch pad end, are 50, 25, and 100 ohms respectively.

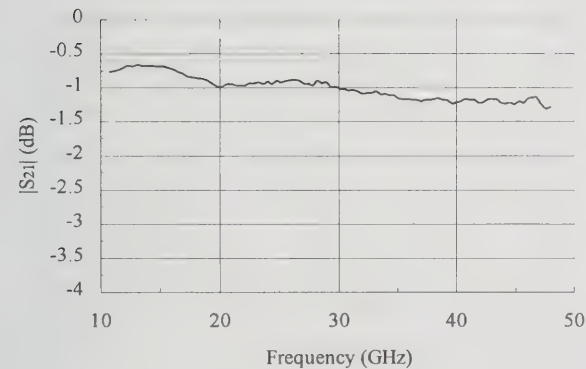


Figure 4. S_{21} for 8mm long, 150 μm deep, transmission line.

An implementation of the impedance variation with line spacing is demonstrated by the construction of a stepped-impedance low-pass filter designed to have a cut-off frequency at 30 GHz. The device is fabricated with six sections: 3 high-impedance (100 ohms) and 3 low-impedance (50 ohms). Figure 4 shows an SEM micrograph of the device. Figure 5 shows the results for S_{21} for the device over the 10-48 GHz frequency range.

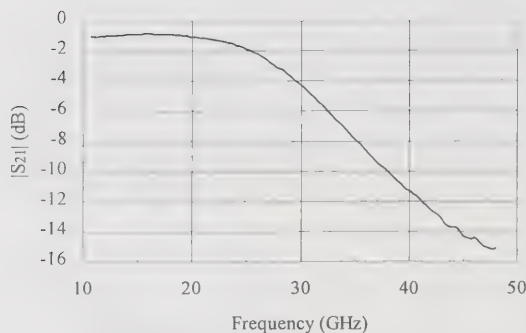


Figure 6. S_{21} for 6-element stepped-impedance filter.

Conclusions

With the achievement of superior attenuation characteristics, the SCREAM process adapted for microwave devices makes possible the fabrication of a whole new set of microwave devices and components. Examples include tunable filters, loaded line phase shifters, variable attenuators, etc. - any device that can utilize the variable impedance of transmission lines or the variable impedance of transmission line stubs. The fact that SCREAM can be performed on fully processed VLSI wafers, and the permission of normal-resistivity wafers which the present modification affords, allows for high compatibility and integrability of these devices with other microelectronics technology.

Acknowledgements

This work is supported by DARPA. All microfabrication was performed at the Cornell Nanofabrication Facility which is supported by the NSF, Cornell University, and Industrial Affiliates. The authors would also like to thank the CNF staff for their assistance. The authors wish to thank Michel Maharbiz for his significant contributions to this project. The authors also wish to thank Noel Elman and Damien Rodger for their great help in the fabrication and testing of the devices.

References

- [1] Veljko Milanovic, Michael Gaitan, Edwin D. Bowen, and Mona E. Zaghoul, "Micromachined microwave transmission line in CMOS technology," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 630-635, May 1997.
- [2] Theodore L. Willke and Steven S. Gearhart, "LIGA Micromachined Planar Transmission Lines and Filters," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp.1681-1688.
- [3] A.A. Ayon, N.C. MacDonald, N.J. Kolias, and R.C. Compton, "Microwave applications of MEMS technology: cantilevered interconnects and tunable transmission lines," 11th International Microwave Conference.
- [4] Arturo Ayon, PhD Thesis, Cornell University, 1996.
- [5] N.C. MacDonald, "SCREAM Micro-Electro-Mechanical Systems," *Microelectronic Engineering*, v. 30, pp.49-73, 1996.
- [6] G.F. Engen and C.A. Hoer, "Thru-reflect-line: An improved technique for calibrating the dual six-port network analyzer," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, pp. 987-993, Dec. 1979.

Resonant Frequency of Dielectric Resonator

Juseop Lee, Sang-Min Han, and Young-Sik Kim, Member, IEEE

Dept. of Radio Sciences and Engineering, Korea University

Anam-dong 5Ga, Sungbuk-gu, Seoul, Korea, 136-701

+82-2-3290-3230, +82-2-921-0544(fax)

E-mail : yskim@kuccnx.korea.ac.kr

Abstract

In this paper, an effective dielectric constant (EDC) has been introduced in an axial attenuation constant of dielectric waveguide model (DWM) method to calculate more accurate resonant frequency of $TE_{01\delta}$ mode of dielectric resonator (DR) in an MIC environment. The accuracy of this method is shown to be better than 1.5 % error.

Introduction

The resonant frequency of a dielectric resonator can be calculated by its dimensions, permittivity, and surroundings [1]. Although techniques predicting the resonant frequency of a dielectric resonator in MIC configuration have been proposed, they have been shown the accuracy of about 1-3.5% due to their assumptions and approximations to simplify the calculations [2], [3]. The existing DWM method that was based on the electromagnetic field continuities at the boundaries on a cylindrical resonator was shown to have an error up to 7% compared with the experimental results [4].

In this paper, an effective dielectric constant has been introduced to predict the resonant frequencies of the $TE_{01\delta}$ mode of a cylindrical and ring dielectric resonator simply and accurately.

Resonant Frequencies

All the field components in each region in Fig. 1 can be obtained from elementary wave functions which should satisfy boundary conditions.

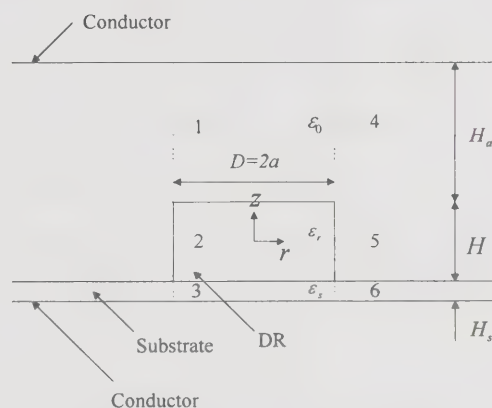


Fig. 1. Cylindrical DR placed in an MIC environment.

From the DWM method, one can derive the following transcendental equation for the radial wavenumber by applying boundary conditions on E_ϕ and H_z at $r = a$ for $-H/2 \leq z \leq H/2$.

$$\frac{J_1(ha)}{hJ_0(ha)} + \frac{K_1(pa)}{pK_0(pa)} = 0 \quad (1)$$

where

J_n : the Bessel function of the 1st kind

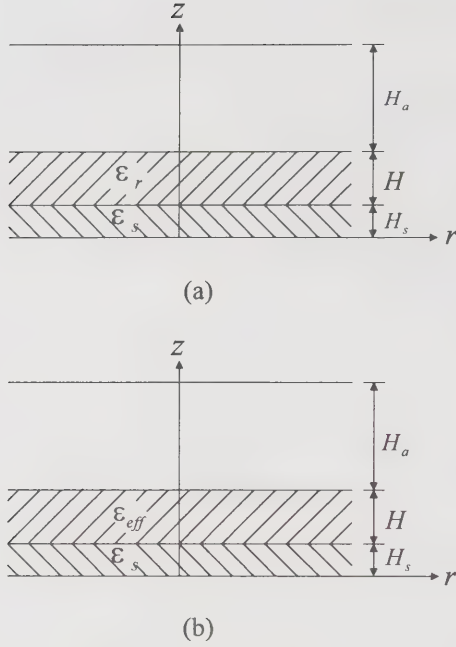


Fig. 2. Analyzed model of (a) DWM and (b) effective dielectric constant method.

K_n : the modified Bessel function of the 2nd kind

$$h^2 = k_0^2 \epsilon_r - \beta^2$$

$$p^2 = \beta^2 - k_0^2 \epsilon_s$$

Assumed the resonator as an infinite radius with same dielectric constant, as shown in Fig. 2(a), the characteristic equation for the axial wavenumber can be given by applying boundary conditions on H_r and E_ϕ at $z = -H/2$ and $z = H/2$ for $r \leq a$.

$$\beta H = \tan^{-1} \left[\frac{\alpha_1}{\beta} \coth(\alpha_1 H_a) \right] + \tan^{-1} \left[\frac{\alpha_3}{\beta} \coth(\alpha_3 H_s) \right] \quad (2)$$

where

$$\alpha_1 = [(\epsilon_r - 1)k_0^2 - \beta^2]^{1/2} \quad (3a)$$

$$\alpha_3 = [(\epsilon_r - \epsilon_s)k_0^2 - \beta^2]^{1/2} \quad (3b)$$

In the case of a ring DR (Fig. 3), following simultaneous equations can be derived for determining

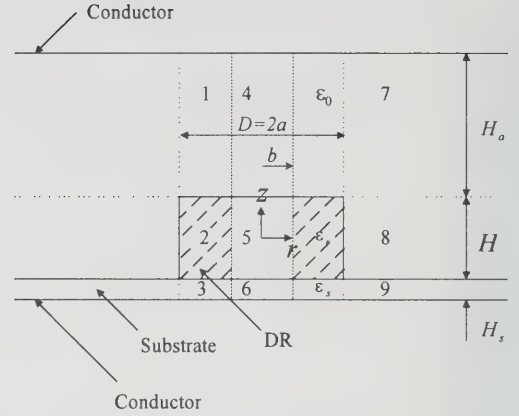


Fig. 3. Ring DR placed in an MIC environment.

the resonant frequency.

$$\frac{I_1(pb)}{qI_0(pb)} - \frac{Z_1(hb)}{hZ_0(hb)} = 0 \quad (4)$$

$$\frac{Z_1(ha)}{hZ_0(ha)} + \frac{K_1(pa)}{pK_0(pa)} = 0 \quad (5)$$

$$\beta H = \tan^{-1} \left[\frac{\alpha_1}{\beta} \coth(\alpha_1 H_a) \right] + \tan^{-1} \left[\frac{\alpha_3}{\beta} \coth(\alpha_3 H_s) \right] \quad (6)$$

where

I_n : the modified Bessel function of the 1st kind

$$Z_n(hr) = J_n(hr) - \frac{J_0(x)}{N_0(x)} N_n(hr)$$

N_n : the Bessel function of the 2nd kind.

In DWM method, one can determine the resonant frequency of a cylindrical DR by solving Eq. (1) and Eq. (2) simultaneously and that of a ring DR by solving simultaneous equations (Eq. (4), (5), and (6)), respectively.

Resonant frequencies can be easily determined by DWM method, but may be resulted in the error of maximum 7% as H_2/a and H_3/a become large. In

the case of microstrip line, the dielectric substrate of finite thickness and a relative permittivity can be replaced equivalently to a homogeneous medium on a hemispheric space with the effective dielectric constant which is smaller than relative permittivity [5]. Similarly, we may introduce the effective dielectric constant of the resonator in the radial direction infinitely, which is different in the axial direction in a microstrip line. We can switch the problem with the structure shown in Fig. 2(a) to that with the structure shown in Fig. 2(b). We define the effective dielectric constant ϵ_{eff} as a weighted average of dielectric constants of a substrate and a dielectric resonator.

$$\epsilon_{eff} = \frac{\epsilon_s H_3 + \epsilon_r H}{H_3 + H}.$$

Using the effective dielectric constant, Eq.(3a) and (3b) are expressed as

$$\alpha_1 = [(\epsilon_{eff} - 1)k_0^2 - \beta^2]^{1/2} \quad (7a)$$

$$\alpha_3 = [(\epsilon_{eff} - \epsilon_s)k_0^2 - \beta^2]^{1/2}. \quad (7b)$$

Using axial attenuation constants of Eq. (7a) and (7b) instead of Eq. (3a) and (3b), one can obtain more accurate resonant frequencies of both cylindrical and ring DRs than DWM method.

Table 1 shows the comparisons of resonant frequencies among experimental result, the DWM method, and this effective dielectric constant method. From Table 1, the effective dielectric constant method yields less than 1.5% in error. Fig. 4, 5, and 6 show the resonant frequency variations as H_a changes. When $H_a = 0.5$ mm, the resonant frequencies obtained by the DWM and the presented methods are almost same and near to the experimental results. As H_a become larger, the resonant frequencies calculated by the presented method are close to the experiment's.

Conclusions

An effective dielectric constant has been introduced in an axial attenuation constant to compute the more

Table 1. Comparisons of resonant frequencies.

ϵ_r	ϵ_s	D (mm)	H (mm)	H_2 (mm)	H_3 (mm)	resonant freq. f_0 (GHz)		
						ref.	DWM [2]	this method
34.19	9.6	14.98	7.48	0.70	0.72	4.35[4]	4.36	4.35
36.2	1.0	4.06	5.15	2.93	2.93	10.50[4]	10.86	10.72
35.2	9.6	4.86	1.81	0.64	∞	12.40[4]	12.89	12.58
36.2	9.5	6.06	4.22	0.70	3.989	8.27[6]	8.44	8.33
36.2	9.5	6.03	3.04	0.70	5.138	9.09[6]	9.33	9.22
36.2	9.5	6.02	2.14	0.70	6.056	10.20[6]	10.51	10.24
36.2	9.5	7.94	2.10	0.70	6.09	8.81[6]	9.13	8.84
34.21	9.6	13.99	6.95	0.70	1.25	4.51[7]	4.53	4.53
34.02	9.6	11.99	5.98	0.70	2.215	5.050[7]	5.09	5.08
36.13	9.6	6.03	4.21	0.70	10.10	8.220[7]	8.48	8.44

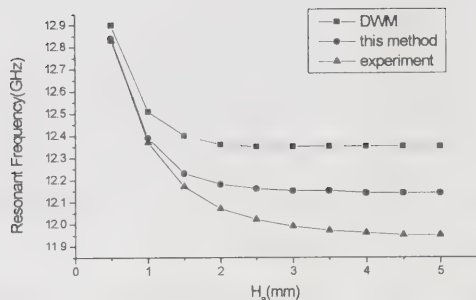


Fig. 4. Resonant frequency of a cylindrical DR.

($\epsilon_r=38.17$, $\epsilon_s=2.2$, $2a=4.63\text{mm}$, $H=2.08\text{mm}$, $H_s=0.508\text{mm}$)

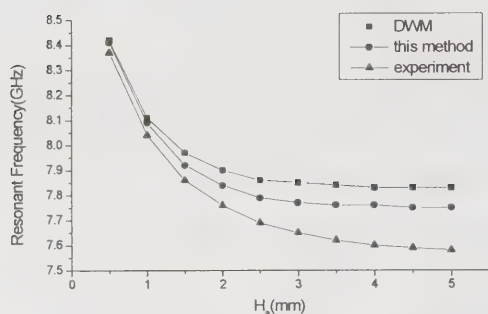


Fig. 5. Resonant frequency of a ring DR.

($\epsilon_r=37.28$, $\epsilon_s=2.2$, $2a=4.63\text{mm}$, $2b=2.0\text{mm}$, $H=3.39\text{mm}$, $H_s=0.508\text{mm}$)

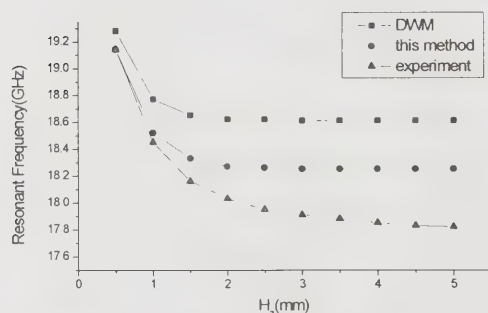


Fig. 6. Resonant frequency of a ring DR.

($\epsilon_r=23.58$, $\epsilon_s=2.2$, $2a=3.94\text{mm}$, $2b=1.3\text{mm}$, $H=1.77\text{mm}$, $H_s=0.508\text{mm}$)

accurate resonant frequencies of cylindrical and ring DRs. It has been shown that the proposed method leads

to much better agreement to the experimental results.

References

- [1] P. Bhartia and I. Bahl, *Microwave Solid State Circuit Design*, John Wiley & Sons, Ch. 3, 1988.
- [2] T. Itoh and R. S. Rudokas, "New method for computing the resonant frequency of dielectric resonators," *IEEE Trans. on Microwave Theory and Techniques*, vol. MTT-25, pp. 52-54, January 1977.
- [3] D. Kajfez and P. Guillon, Editors, *Dielectric Resonators*, Artech House, Ch. 1, 1986.
- [4] R. K. Mongia, "Resonant frequency of cylindrical dielectric resonator placed in a MIC environment," *IEEE Trans. on Microwave Theory and Techniques*, vol. MTT-18, pp. 802-804, June 1990.
- [5] D. M. Pozar, *Microwave Engineering*, Addison-Wesley, Ch. 4, 1993.
- [6] M. W. Pospieszalski, "Cylindrical dielectric resonators and their applications in TEM line microwave circuits," *IEEE Trans. on Microwave Theory and Techniques*, vol. MTT-27, pp. 233-238, March 1979.
- [7] R. K. Mongia and P. Bhartia, "Accurate conductor Q-factor of dielectric resonator placed in an MIC environment," *IEEE Trans. on Microwave Theory and Techniques*, vol. MTT-41, pp. 445-449, March 1993.

Design of an Integrated RF Filter for the Direct Digitization Front End of a Dual GPS/GLONASS Software Radio Receiver

Cornelius Brinegar and Krishna Naishadham

Department of Electrical Engineering

Wright State University

Dayton, OH 45435

email: knaish@cs.wright.edu

Tel. (937)775-5057

Fax: (937)775-5009

Abstract

A software radio can be defined as a receiver implementation in which the analog-to-digital (ADC) converter is positioned as close to the antenna as possible, so that the resulting samples can be processed using a programmable microprocessor. Compared to the conventional superheterodyne implementation, this direct digitization front-end eliminates additional components used for frequency translation. A key component of such a front-end is the RF filter which facilitates band-pass sampling of the incoming signal. In this paper, we present the design of a wide-band integrated filter which passes both the GPS and GLONASS satellite bands using a minimum number of analog components. The filter is realized on a microstrip and simulated to test the baseband specifications and signal integrity.

1 Introduction

A significant advance in receiver design since the advent of superheterodyne principle in 1918 is the development of the software radio concept [1]. The traditional superhet front end consists of multiple stages of frequency down-conversion and analog signal amplification, followed by an ADC (digital processing) or a crystal detector (analog processing). The goal in developing any software radio is to place the ADC as close to the antenna as possible, and process the resulting samples on a programmable microprocessor, without the need for additional components for frequency down-conversion. The front end of a software radio is typically referred to as direct digitization front end (DDFE). The elimination of analog components for frequency translation in the DDFE has many advantages [2]: (a) the potential sources of error introduced by the non-linearity of the mixing and demodulation process, such as total harmonic distortion, can be avoided, (b) the processing of digital samples by a programmable microprocessor offers flexibility in evaluating a particular receiver architecture or the incorporation of the latest signal processing algorithms without the need to build costly prototypes, (c) because the ADC in

software radio operates at high conversion rates, thus capturing a broad frequency spectrum, a single front end can be used to process a wide range of RF transmissions. Thus, a single software radio has the potential to replace multiple receivers.

The system level considerations and design trade-offs for software radio have been described in [3], emphasizing a novel technique of band-pass sampling of multiple signals, which allows for the processing of both GPS and GLONASS signals. Bandpass sampling or intentional aliasing can fold multiple frequency bands (such as GPS and GLONASS) into a single baseband by sampling the incoming signal directly at the carrier frequency. In this approach, the DDFE consists of an antenna, a pre-amplifier, a high-speed ADC, and a programmable microprocessor at the least [3]. However, noise and dynamic range considerations demand additional amplification stages to obtain the necessary gain, and bandpass filtering to suppress unwanted out-of-band radio interference. To process GPS/GLONASS signals simultaneously, Akos and Braasch [2] adopted the standard approach of using a power divider after the pre-amplification stage to split the received signal into GPS and GLONASS channels. Then, one

branch would filter the GPS signal while the other processes the GLONASS channel. Next, outputs from these separate channels would be recombined to make the two bands available to the ADC.

An alternate solution using a fewer number of components is proposed in this paper, namely, to create a filter system with effectively two pass bands, one for GPS and one for GLONASS. Such a filter is composed of only two analog filters which are cascaded in series. The first, a bandpass filter, would pass both the GPS and GLONASS bands. The second, a bandstop filter, would attenuate the out-of-band frequencies between the GPS and GLONASS bands. Advantages of this system include only two main components, potentially small size, better signal-to-noise (S/N) ratio by excluding the power divider. In fact, the pre-amplification, although not subject of the present paper, can be accomplished more easily than in the traditional approach, because the S/N ratio of each stage can be relaxed. Thus, by using the fewest number of analog components in the system, signal integrity is improved throughout the analog path. The design and simulation of the integrated GPS/GLONASS filter is described next.

2 Filter Transfer Characteristics

The first step in the design was to determine the response of the filters. This means finding the high and low cutoff frequencies as well as the roll-off rate. Figure 1 is a block diagram illustrating the specifications for the two filters. Also, it was decided to use maximally flat characteristic, as this type of filter can achieve the required attenuation with a reasonable order, provide a fairly linear phase delay, and be smooth in the pass band. To meet these specifications (Figure 1), it was calculated from the Filter Handbook [4] that the bandpass filter will need to be of fifth order, and the bandstop filter must be of fourth order.

The filters can be specified in terms of a typical ladder circuit using inductors and capacitors. The lowpass prototype elements and transformation formulas are well-known (*cf.* [5]), and were used to calculate the component values of the ladder circuit representation of the filters. Figures 2 and 3 show the resulting transfer function re-

sponse of the bandpass and bandstop circuits, respectively, calculated from the ladder circuit representations. For both transfer function plots, 6 dB was added to account for the loss due to voltage division between the source and load resistances. The graphs correspond extremely well with the specifications drawn in Figure 1.

3 Implementation Method

Lumped element filters will not be used to realize the design. This means that the ladder circuits will have to be approximated by transmission line circuits. To accomplish this task with minimum number of components, it was decided to design the bandpass filter using coupled lines, and the bandstop filter using L-C resonators. Most other filter types require unrealizable impedance values because of the fractional bandwidth of the filters being on the order of 1%.

The actual realization of the filters was an involved process. For the bandpass filter, the following design equations [4], [5] allow for calculation of the even and odd impedances for each coupled section:

$$Z_0 J_1 = \sqrt{\frac{\pi \Delta}{2g_1}} \quad (1)$$

$$Z_0 J_n = \frac{\pi \Delta}{2\sqrt{g_{n-1}g_n}}, \quad n = 2, 3, \dots, N \quad (2)$$

$$Z_0 J_{N+1} = \sqrt{\frac{\pi \Delta}{2g_N g_{N+1}}} \quad (3)$$

where Z_0 is the characteristic impedance of the input line and J_n is a normalized constant for the n^{th} quarter-wave admittance inverter. The even and odd mode impedances follow from

$$Z_{0e} = Z_0(1 + JZ_0 + (JZ_0)^2) \quad (4)$$

$$Z_{0o} = Z_0(1 + JZ_0 + (JZ_0)^2). \quad (5)$$

The calculated values for the bandpass filter are shown in Table 1, and the coupled-line filter layout was realized on a microstrip substrate ($\epsilon_r = 10.2$ (Duroid 6010), thickness $H = 1.27$ mm) using these design constants. The choice of this substrate was facilitated by the realization of the range of required even and odd mode impedances (see Table 1) using a normalized strip width close

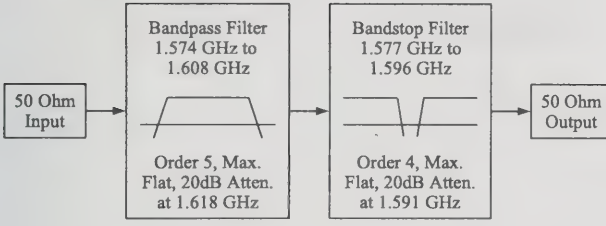


Figure 1: Block Diagram of System Design with Specifications

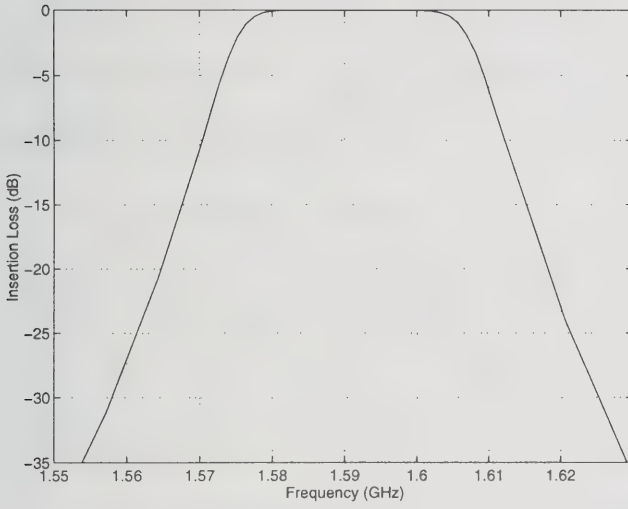


Figure 2: Spice Response of the Bandpass Filter

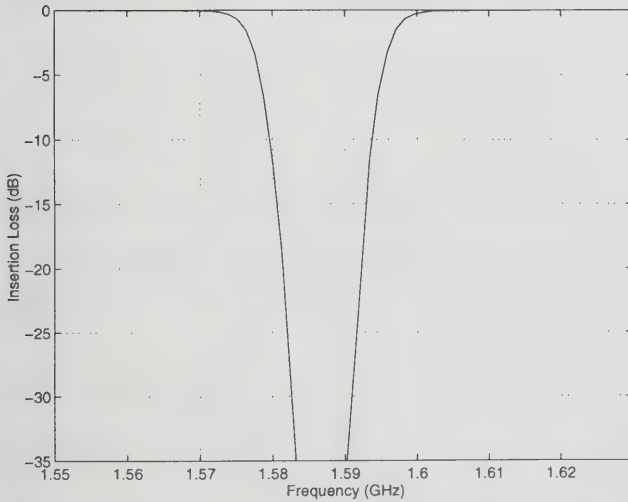


Figure 3: Spice Response of the Bandstop Filter

n	g_n	$Z_0 J_n$	$Z_{0e} (\Omega)$	$Z_{0o} (\Omega)$
1	0.6180	0.23160	64.26	41.10
2	1.6180	0.03315	51.71	48.40
3	2.0000	0.01843	50.94	49.10
4	1.6180	0.01843	50.94	49.10
5	0.6180	0.03315	51.71	48.40
6	1.0000	0.23160	64.26	41.10

Table 1: Design Constants of the Coupled-Line Bandpass Filter.

n	g_n	$\chi_n (\Omega)$	$\phi_n (\text{deg})$	$C_n (\text{pF})$
1	0.7654	5579	173.2	0.239
2	1.8478	2311	169.4	0.375
3	1.8478	2311	169.4	0.375
4	0.7654	5579	173.2	0.239

Table 2: Design Constants of the Bandstop Filter.

to $W/H = 1$ and the normalized spacing close to $S/H = 2.5$ to 3. These dimensions turned out to be within fabrication tolerances.

The bandstop filter's layout was determined from the design equations and tables in the Filter Handbook [4]. We first assume that all the lines in this filter have a characteristic impedance of 50Ω . This value was fixed to assure that the line widths could be achieved easily. Then the reactance slope parameter for each section (χ_n), the electrical length of the open resonator (ϕ_n), and the gap capacitance (C_n) for each section, was calculated from standard equations to yield the design shown in Table 2. We could achieve compact placement of the resonator sections with the desired attenuation only by using lumped chip capacitors in place of capacitive gaps.

4 Simulation

To determine the validity of the designs before construction began, the two circuits were simulated. The bandpass filter was simulated in IE3D, and its S-parameters were found. At first it was noticed that the pass band was not placed at quite the right frequency, so the resonators were shorted to shift the response up. A final length correction of -0.2 mm was determined to give the cor-

rect placement of the pass band. The simulated response of the bandpass filter showed the cut-offs and roll-offs that were expected. The magnitude response was reasonably flat and the phase shift linear over the band from 1.57 GHz to 1.61 GHz. A slight shift of about -0.5 dB was observed relative to the lumped circuit prototype response, because of parasitic-induced losses such as radiation.

5 Summary

In summary, this paper dealt with the design of two filters which can be cascaded to replace a four component system in the front end of a software radio for receiving both the GPS and GLONASS bands. Undesirable losses in the power divider and separate bandpass filters for the two frequency channels have been avoided. Both the filters have been simulated and shown to have the desired characteristics. The next step of the process will be to fabricate the two filters and integrate them on a single mother-board with all the connectors and lumped components. We anticipate to present the comprehensive design details of this system at the conference along with experimental corroboration.

References

- [1] *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26-68, May 1995.
- [2] D.M. Akos and M.S. Braasch, "A Software Radio Approach to Global Navigation Satellite System Receiver Design," *Proc. Institute of Navigation Annual Mtg.*, Boston, MA, June 1996.
- [3] D.M. Akos and J.B.Y. Tsui, "Design and Implementation of a Direct Digitization GPS Receiver Front End," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-44, pp. 2334-2339, Dec. 1996.
- [4] G. Matthaei, L. Young, and E.M.T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, Dedham, MA: Artech House, pp. 725-774, 1980.
- [5] D.M. Pozar, *Microwave Engineering*, Addison-Wesley, New York, 1990, pp. 455-528.

MODELING AND CHARACTERIZATION OF WIRE BONDING AND TAPED AUTOMATIC BONDING. ANALYTICAL FORMULAS AND EXPERIMENTAL VALIDATION

N. Hassaine, Y. Shen and P. Ntake

Harris Farinon Canada, Inc.

3 Hotel de ville, Dollard-Des-Ormeaux

Qc, Canada, H9B 3G4

Tél: 1 (514) 421 8447, Fax: 1 (514) 421 3756

e-mail: NHassain@Harris.com

Topic:

Passive components: packaging technology

Abstract:

A simplified method has been established to model and characterize two types of short interconnections: wire bonding and Tape Automatic Bonding (TAB). A comparison between these two parasitic elements is presented and analytical formulas easy to use in CAD applications are proposed to calculate their inductance. These results are compared with measurements performed in the frequency range of (1-30) GHz.

I. INTRODUCTION:

The continuous improvement of integrated circuit fabrication process and the increasing circuit complexity (MMICs and MHMICs) allow to increase the integration density and give rise to increasing number of inputs and outputs. At high levels of integrations, a large number of short interconnections is needed which induces delays, interference and disturbs the signal propagations. For these reasons, circuit simulators should take into account parasitic effects. The present work is an attempt to study two types of these parasitic elements: wire bonding and Tape Automatic Bonding (TAB) figure 1.

II. FORMULATION OF THE PROBLEM:

Bonding wires and TAB are generally described by an inductance which value is determined from the potential vector integral [1]. Inductances are considered for a general N-loop system, which is representative of the conductor arrangements of interest. Assuming that the current direction is given by the axis of the conductor, the potential vector is expressed by the known relation:

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \sum_{j=0}^{N_c} \int_{vol_j} \frac{J(\vec{r}_j)}{|\vec{r} - \vec{r}_j|} dS_j \vec{dl}_j \quad (1)$$

where $J(\vec{r}_j)$ is the current density at the point \vec{r}_j , inside the elementary volume $dvol_j$ of length dl_j and cross-section dS_j of the conductor j and where N_c is the total number of sections forming the conductor.

If we assume that a segment of conductor i is maintained at a constant voltage, the overall voltage drop across this conductor on a cross section S_i can be expressed by:

$$V_i = j\omega \sum_{j=0}^{N_c} \iint_{S_i l_i} \vec{A}_{ij} \vec{dl}_i dS_i + \iint_{S_i l_i} \frac{J(r_i)}{\sigma_i} dl_i dS_i \quad (2)$$

which can be written as :

$$V_i = j\omega \sum_{j=0}^{N_c} L_{p_{ij}} + R_{p_i} \quad (3)$$

where :

$L_{p_{ii}}$ is the partial self inductance of conductor i ,

$L_{p_{ij}}$ is the partial mutual inductance between conductor i and conductor j ,

R_{p_i} is the partial resistance of conductor i .

The coefficients are difficult to evaluate analytically from the above integrals, since the current distribution is generally unknown and the shapes of the conductors variables. We have thus used the Partial Element Equivalent Circuit method (P.E.E.C.) [2], which consists in dividing the conductors into elementary cylinders where uniform current is assumed. Each elementary cylinder is represented by its partial resistance, its partial self and mutual inductances.

III. ANALYTICAL FORMULAS:

Electromagnetic simulation techniques require large computer memories, long computing time and large data handling. Thus, they are difficult to use extensively in CAD systems. We have developed analytical formulas to calculate the inductance of the wire bonding and the TAB cases. We derive them from extracted results using the least square method [3]. Because of their empirical nature, these expressions are restricted to a specified domain of validity.

III. 1- One wire bonding

The considered geometry consists of a wire bonding with a circular profile of radius R and diameter d (figure 1a). In this case, the analytical formulas have an accuracy of about 2.5% for the following set of parameters:

$$10 \leq d(\mu m) \leq 70$$

$$100 \leq R(\mu m) \leq 650$$

The detailed expression for this case is given below:

$$L(nH) = a(R^2) + \left(\sum_{i=1}^4 b_i d^{4-i} \right) (R) + \left(\sum_{i=1}^4 c_i d^{4-i} \right) \quad (4)$$

where $a = 9.6786 \cdot 10^{-7}$ and Table 1 lists the coefficients b_i and c_i ($i = 1$ to 4).

	$i=1$	$i=2$	$i=3$	$i=4$
b_i	$-4.630 \cdot 10^{-9}$	$7.540 \cdot 10^{-7}$	$-5.485 \cdot 10^{-5}$	0.0032
c_i	$3.241 \cdot 10^{-8}$	$-6.528 \cdot 10^{-6}$	$8.829 \cdot 10^{-4}$	-0.0835

Table 1: Coefficients b_i and c_i for the case of one wire bonding.

III.2- Ribbon connection

This mode of connection (TAB: Tape Automatic Bonding) is generally more performant than wire bonding because it reduces the value of the inductance. The considered geometry consists in a ribbon of width W , length l and height H to the ground plane (figure 1b). For the following set of parameters:

$$10 \leq W(\mu m) \leq 150$$

$$80 \leq H(\mu m) \leq 500$$

$$100 \leq l(\mu m) \leq 600$$

The closed form expression for this case is given below:

$$L(nH) = A \ln(W) + B \quad (5)$$

where

$$A = 0,0028 - (1,621 \cdot 10^{-4})l$$

$$B = B_1 \times (LnH)^2 + B_2 \times (LnH) + B_3$$

$$B_1 = B_{11} \times (l)^2 + B_{12} \times (l) + B_{13}$$

$$B_2 = B_{21} \times (l)^2 + B_{22} \times (l) + B_{23}$$

$$B_3 = B_{31} \times (l)^2 + B_{32} \times (l) + B_{33}$$

Table 2 lists the coefficients B_{ij} ($i=1$ to 3 and $j=1$ to 3).

		i=1	i=2	i=3
B_{ij}	j=1	$1.125 \cdot 10^{-8}$	$3.670 \cdot 10^{-5}$	0.0026
	j=2	$-1.339 \cdot 10^{-8}$	$4.347 \cdot 10^{-4}$	-0.0313
	j=3	$1.351 \cdot 10^{-9}$	$1.522 \cdot 10^{-6}$	$3.112 \cdot 10^{-4}$

Table 2: Coefficients B_{ij} for the inductance of a ribbon connection.

IV. COMPARISON WITH SIMULATIONS AND EXPERIMENTAL RESULTS

Experimental measurements, performed at the France Telecom CNET Bagneux Laboratory, have been obtained on various test boards in the frequency range 1-30 GHz. Simulations have been carried out with HFSS (High Frequency Structure Simulator). A good agreement (within 4%) is obtained with electromagnetic simulations as well as with experimental results. In figure 3, analytical formulas are compared with the theoretical results obtained by numerical modeling for the case of one wire bonding and ribbon connection. Measured and simulated S_{11} and S_{12} parameters for the two cases are plotted on figure 4.

V. CONCLUSION:

A simplified method to calculate the inductance of wire bonding and Tape Automatic Bonding (TAB) has been set up which allows to take into account precisely and easily the parastic effects in circuit design. Analytical formulas easy to use in CAD are derived from the numerical results using the least square method. The formulas have been shown to agree, with a precision in the order of 4%, with both simulation and experimental results in the frequency range of (1-30) GHz

VI. REFERENCES:

- [1] *A. Ruehli, P. A. Brennan*, "Capacitance models for integrated circuit metallization wires" IEEE journal of solid-state circuits, sc-10, n°6; Dec. 1975.
- [2] *R. B. Wu, C. N. Kuo and K. K. Chang*, " Inductance and resistance computations for three-dimensional multiconductor interconnection structures", IEEE Trans. Microwave Theory Tech., vol. 40, pp. 263-271, Feb. 1992.
- [3] *N. Hassaine*, "Modélisation des éléments passifs de circuits intégrés GaAs. Applications à la logique ultrarapide, à l'optoélectronique et aux hyperfréquences", Thèse de Doctorat (Ph. D. dissertation), INP Grenoble, France, Feb. 1994.

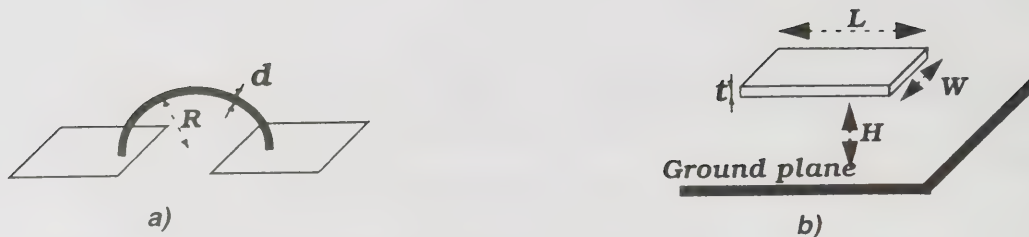
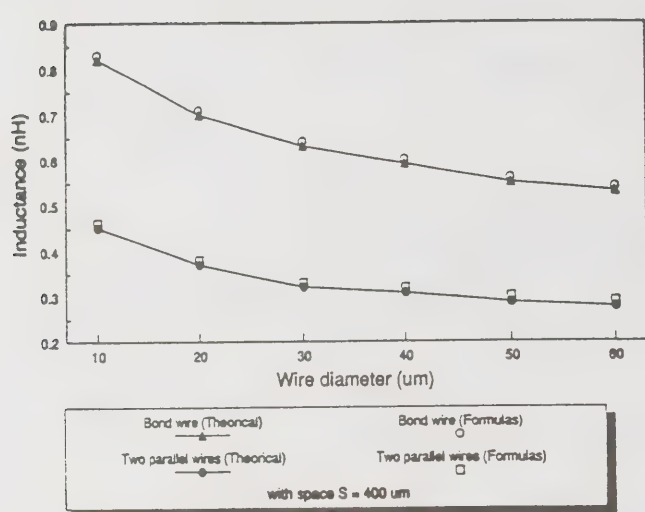
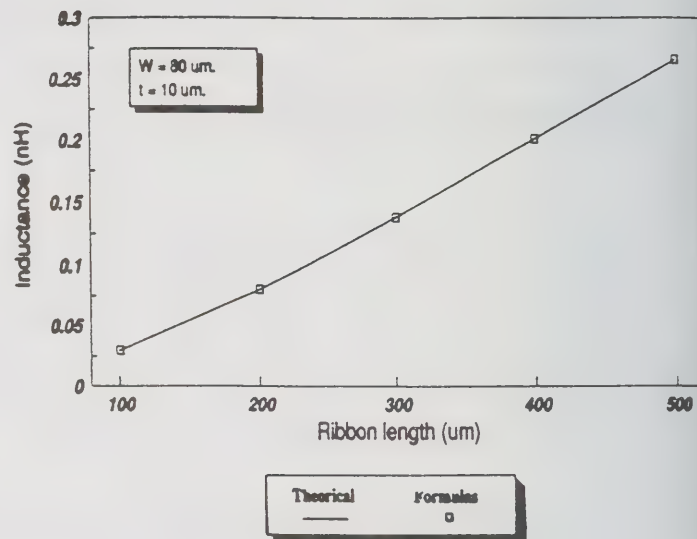


Figure 1: Different types of interconnections. a)- Wire bonding. b)- TAB connection.



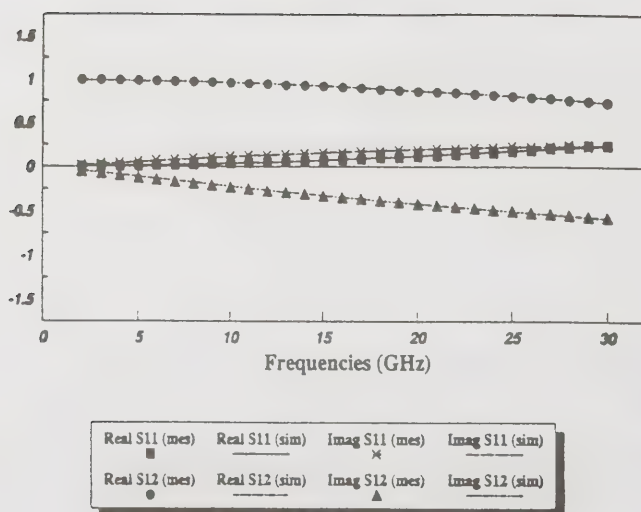
a)



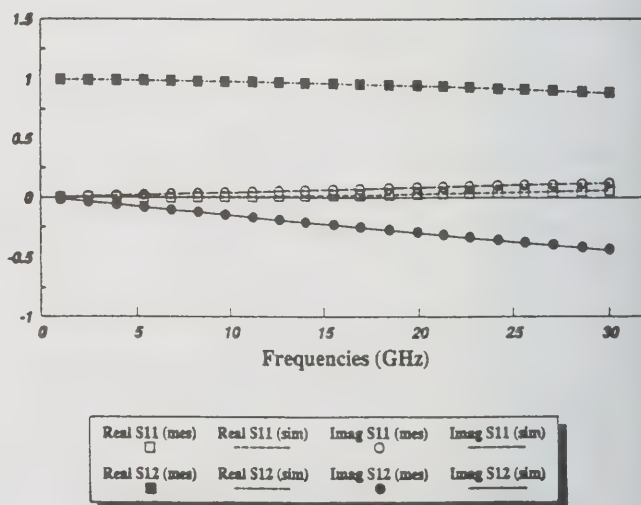
b)

Figure 2: Comparison between analytical formulas and numerical results.

a)- Wire bonding. b)- TAB



a)



b)

Figure 3: Measured and simulated S_{11} and S_{12} parameters (1-30) GHz. a)- Wire bonding. b)- TAB

Advanced Silicon IC Interconnect Technology: Present Trends and RF Wireless Implications

Ronald J. Gutmann

Center for Integrated Electronics and Electronics Manufacturing
Rensselaer Polytechnic Institute
Troy, NY 12180-3590
518-276-6794, 518-276-8761 (fax), rgutmann@unix.cie.rpi.edu

Abstract

Back-end-of-the-line (BEOL) trends in silicon ICs include fully planarized interconnect structures with six levels on non-local wiring, copper metallization for improved resistance and electromigration and low dielectric constant (low-k) interlevel dielectrics (ILDs) for reduced line and coupling capacitance. These technological advances, when combined with front-end silicon technology innovations, will impact technology trends in RF wireless ICs and enhance the potential for systems-on-a-chip and application specific ICs (ASICs) with embedded communications capability. Design complexity will be alleviated by use of intellectual property (IP) cores and virtual design environment (VDE) software. Si IC BEOL trends are summarized, synergistic front-end developments discussed, implications for wireless technologies presented, the impacts of a VDE with IP cores discussed, and a timetable for practical realization in wireless products projected.

Key words: Si IC interconnect technology, wireless technology trends, embedded communications, intellectual property cores

Introduction

Wireless communication technology that is highly leveraged to, and compatible with, mainstream Si CMOS ICs will have a significant cost advantage and will benefit from accelerated technology performance enhancements driven by logic and high-volume application specific ICs (ASICs). In the past decade, significant changes in Si IC technology have taken place in the back-end-of-the-line (BEOL) which have been/are being/will soon be incorporated in mainstream Si IC products; these changes benefit monolithically integrated Si RF circuits, enable embedded wireless communications and enhance prospects for systems-on-a-chip.

The BEOL technology advances are mainly three-fold: chemical-mechanical planarization (CMP) that enables 5 to 8 levels of on-chip metallization, copper metallization to replace aluminum wires and tungsten vias for improved electrical resistance and electromigration, and lower dielectric constant (low-k) interlevel dielectrics (ILDs) to replace SiO₂ for lower line and coupling capacitance. These three Si IC technology changes which are well established in the last five years, are being introduced in mainstream manufacturing and will be introduced in manufacturing in the next 2 to 4 years, respectively, enhance RF passive circuit performance.

An unrelated development, synergistic for embedded RF wireless technologies, is occurring in Si foundries and IC ASIC design. Because of increased design complexity and inherent inefficiencies in conventional gate array and standard cell implementations, fully functional proprietary macrocells are becoming building blocks for ASICs; these macrocells are often referred to as intellectual property cores (IP cores). In parallel, the complexity of IC design of microprocessors and complex ASICs has resulted in apriori partitioning of the design, floor planning of major building blocks and increasingly large design teams to meet decreasing product design cycles with first-pass (or second-pass) success. A virtual design environment (VDE) compatible with large spatially-distributed design teams is becoming increasingly required, in spite of the firewalls desired to protect proprietary information. In a more agile business environment, firewalls will be reduced over time as leading edge manufacturing capabilities are more readily accessible (the capability gap between Si foundries and leading-edge vertically integrated companies continues to decrease).

This paper summarizes Si BEOL trends, namely CMP, copper metallization and low-k ILDs, briefly presents synergistic front-end developments (reduced feature size, Si-Ge technology and high-k materials for

storage capacitors and MOS gates), presents RF advanced wireless technology projected with these developments, discusses briefly the important synergistic trends in VDE and IP cores, and projects a timetable for such implementations compatible with the National Technology Roadmap for Semiconductors (NTRS).

Silicon IC Interconnect Trends

The three main BEOL innovations that affect interconnect limited Si ICs, specifically microprocessors and ASICs, are CMP, copper metallization and low-k ILDs. All are the result of intense research and development efforts to reduce the impact of interconnection on electrical performance and yield as minimum feature size has migrated from 750 nm to 180 nm in the past decade, with 100 nm technology of current research interest.

Global planarization using CMP was the first of these innovations, developed by IBM and used in IC products first by IBM, Intel and Micron within the last decade. Presently, all major IC manufacturers use CMP for microprocessors and ASICs, where 4 to 6 levels of metallization are required. The planarization provided by CMP allows an almost unlimited stacking of metallization levels in ICs, limited by yield considerations and real estate requirements for conducting vias to interconnect different levels (and the Si devices) [1]. The CMP equipment and consumable business has been the most rapidly growing part of the IC equipment industry with approximately 30% compound annual growth rate (CAGR) in the past few years.

Copper metallization has been actively pursued in many laboratories for a decade, as improved electrical conductivity and electromigration capability can be achieved [2]. The announcement in September 1997 by IBM that copper will be used in ICs in the latter half of 1998 has intensified such efforts. The IBM six-level copper capability (with oxide ILDs and tungsten local interconnects) not only has superior electrical performance (clock speed in comparably design RISC processors and improved electromigration capability compared to aluminum alloys), but is claimed to be a lower-cost higher-yield manufacturing process. Ta, TaN, and TiN are the liners of choice, with electroplating of Cu on a Cu seed layer deposited over the liner the process of choice at this time. Copper will be used in microprocessors and high performance ASICs on a broad scale in the next few years. The equipment and consumable suppliers have expanding development programs, similar to CMP expansions in the early '90s.

Since reactive ion etching (RIE) of copper is not possible without elevated processing temperature, dual damascene patterning of copper is required. In this process, vias and trenches are etched into a planarized ILD and copper (and the required liners) are deposited. CMP is used to remove the excess copper in the trench regions and over the dielectric, as illustrated in Figure 1 [3]. The damascene patterning process is a major processing change that needs to be incorporated with copper; however, damascene patterning eases the fabrication of three-dimensional passive structures for RF circuits.

The third innovation is the use of low-k materials compatible with copper metallization [4,5]. While appreciable research and development is underway at SEMATECH, industrial laboratories and universities, no product and manufacturing announcements have been made at this time. Since some of the low-k research has emphasized the use in aluminum interconnects (where gap fill and planarization are more important than RIE profiles and Cu-damascene compatibility), future directions are less clear. However, the performance leverage of low-k ILDs with Cu metallization is significant for microprocessors and ASICs; product announcements can be anticipated within a few years.

Synergistic Silicon IC Front-End Developments

While the BEOL developments described have great performance leverage for RF wireless products, Si front end technologies compatible with mainstream Si IC fabrication need be considered. These include continual reduction in minimum feature size (MFS), use of Si-Ge heterojunction technology, and alternative high-k dielectrics for storage capacitors and MOSFET gates.

Continued developments in lithography and pattern transfer have resulted in optical stepper lithography achieving 250 nm MFS in manufacturing today, with 180 nm demonstrated in various laboratories; while the lithography of choice at 100 nm is unclear (130 nm will probably be achieved with optical), various choices indicate the 100 nm and 80 nm will be achieved compatible with the NTRS. Scaling-induced performance enhancements for Si ICs on 8" and 12" diameter wafers are tremendous leverages for process-compatible RF wireless implementations.

The use of Si-Ge heterojunction bipolar transistor (HBT) technology has been demonstrated to be compatible with high volume Si IC manufacturing by IBM [6]. While additional hetero-epitaxial layer growth is necessary, the wafer processing can be done on a

high-performance manufacturing facility in a manner compatible with high performance MOSFETs. Clearly high performance IP cores of RF components using Si-Ge HBTs can be embedded in high performance MOSFET-based ICs in a monolithic fashion.

Innovative high-k thin film dielectrics are being developed for reduced-area storage capacitors for memory devices and for alternative gate dielectrics for scaled CMOS technologies [5]. While these ferroelectric materials (eg. titanates) require high temperature deposition to obtain highest dielectric constant, deposition at lower temperatures compatible with BEOL processing provides a potentially useful dielectric constant (~30) which can be useful in RF filtering applications and on-chip power decoupling.

Wireless Technology Implications

The availability of 5 to 8 levels of planarized, damascene-patterned copper metallization in the next 1 to 2 years combined with the addition of low-k dielectrics a few years later, offers improved RF passive components on-chip [7]. In most monolithic implementations, inductor Q is a serious performance limiter, and the low Q is diminished further with Si substrate conductivity. Multilevel interconnects allow the implementation of multiturn vertically oriented toroidal inductors, with increased inductance per unit chip area and increased Q; processing of such structures is simplified with dual damascene patterning. Well grounded vias can be connected to ground (or bypassed power) planes near the top levels of the on-chip interconnect structure to maximize RF performance. Closely coupled toroidal inductors are feasible, resulting in many lumped element alternatives for conventional RF/microwave components (such as high Q lumped element realizations of quadrature hybrids).

In addition, transmission line interconnects can be realized in either a coaxial configuration (using four metal levels), a stripline configuration (three levels), a microstrip configuration (two levels) or a coplanar waveguide (potentially one level). Such transmission lines can interconnect RF macrocells located in different areas of an ASIC, allowing chip partitioning to minimize RF macrocell interference and allowing a small density of distributed circuit components as necessary.

Various lumped antenna elements can also be implemented on-chip, in particular structures like patch antennas. With multilevel capability, the top level can be entirely devoted to the antenna, increasing the radiation efficiency compared to smaller structures.

More effort in integrated antenna structures seems warranted.

With RF grounding through the interconnect structure, thinned Si ICs are feasible to decrease thermal resistance (thermal path through the substrate). Signals and grounds are brought to the top surface of the chip. Flip chip mounting is clearly the direction of packaging future logic and ASICs with an increasing number of pin outs required [8].

Virtual Design Environment with IP Cores

The VDE envisioned includes compatible CAD tools with graphical interfaces, standard neutral data files and a common object request broker architecture (CORBA) or equivalent within a proper information network (such as Valor's Enterprise 3000) [9]. Such a system would include electrical CAD, floor planning and layout capability; macrocells and/or IP cores will be an integral part of the data base. Such a capability is already demonstrated for printed circuit board implementations; migration to the IC level will be a tremendous enabler for rapid turn around ASICs with manufacturing robustness and first-pass design success. Such an environment will require further research and development as well as data base and networking standards. However, the increasing IC design complexity and economic leverage from such a capability are significant drivers for research and near-future implementation.

IP cores of interest in this context include performance-driven, technology-specific, product-differentiating macrocells (so-called hard IP) such as microprocessors, digital signal processors (DSPs), mixed signal blocks and embedded RF wireless cells [10]. The main feature is the reuse capability and apriori specified functionality of such process-dependent macrocells. Libraries have been established that are compatible with multiple foundries using mainstream CMOS technology; however, the migration to copper metallization and low-k ILDs will result initially in foundry-dependent RF IP cores, particularly when these BEOL innovations are combined with front-end developments like Si-Ge HBTs and high-k dielectrics. The long-range trend envisioned is for the Si chip to be analogous to a motherboard, with the IC design analogous to PCB design today.

Timetable Projections and Conclusions

The three BEOL innovations, the front-end developments and the trend toward a VDE with core-based designs for system-on-a-chip (SOC) are projected in the recent version of the NTRS [11]. They are

projected for full implementation before 2006 at the 100 nm technology node.

During this full-implementation period, the trend toward Si monolithic implementation of wireless technology products will accelerate. The economic and performance pace of technological innovation will outstrip improvements in compound semiconductor technology. Three-dimensional implementations could result in an increasing role for compound semiconductor technology, particularly if the pace of Si IC technology slows after 100 nm MFS. Compound semiconductor front ends will remain dominant at higher microwave frequencies and in hybrid implementations. However, the trends described indicate that Si ICs will increasingly be the solution of choice whenever performance objectives can be achieved.

Acknowledgements

This work has been supported by the SEMATECH Center of Excellence (SCOE) in Multilevel Interconnects, the SRC Center for Advanced Interconnect Science and Technology (CAIST) and the DARPA/NSF Electronics Agile Manufacturing Research Institute (EAMRI) at Rensselaer. Interactions with Profs. Robert J. Graves and Shyam P. Murarka have assisted in shaping these perspectives and are gratefully acknowledged.

References

1. J.E. Steigerwald, S.P. Murarka and R.J. Gutmann, *Chemical-Mechanical Planarization of Microelectronic Materials*, J. Wiley and Sons, 1997.
2. D. Edelstein, et al., "Full Copper Wiring in a sub-0.25 μ m CMOS ULSI Technology", International

Electron Devices Meeting, 1997, pp.773-776.

3. R.J. Gutmann, D.T. Price, J.M. Neiryneck, C. Saino, D. Permana, D.J. Duquette and S.P. Murarka, "CMP of Copper-Polymer Interconnect Structures", CMP-MIC Conference, 1998, pp.257-266, invited.

4. R.J. Gutmann, T.P. Chow, D.J. Duquette, T.M. Lu, J.F. McDonald and S.P. Murarka, "Low Dielectric Constant Polymers for On-Chip Interlevel Dielectrics with Copper Metallization", Mat. Res. Soc. Proc., Vol. 381, 1995, pp. 177-195, invited.

5. R.J. Gutmann, W.N. Gill, T.M. Lu, J.F. McDonald, S.P. Murarka and E.J. Rymaszewski, "Low- and High-Dielectric Constant Thin Films for Integrated Circuit Applications", Advanced Metallization and Interconnect Systems for ULSI Applications in 1996, Mat. Res. Soc., 1997, pp. 393-400, invited.

6. B. Myerson, "UHV/CVD Growth of Si and SiGe Alloys: Chemistry, Physics and Device Applications", Proc. IEEE, **80**, 1992, pp. 1592-1608.

7. A.C. Reyes, S.M. El-Ghazaly, S.J. Dorn, M. Dydyk, S.K. Schroder and H. Patterson, "Coplanar Waveguides and Microwave Inductors on Silicon Substrates", IEEE Trans. Microwave Theory Tech, **43**, 1995, pp. 2016-2022.

8. R.R. Tummala and E.J. Rymaszewski, *Microelectronics Packaging Handbook*, Chapter 6, Van Nostrand Reinhold, 1989.

9. R. Subbu, A.C. Sanderson, C. Hocaoglu and R.J. Graves, "Distributed Virtual Design Environment Using Intelligent Agent Architecture", Proceedings of the Industrial Engineering Research Conference, May, 1998.

10. P. Brown and A. Stefora, "IP: The Road Doesn't Have To Be Hard", Electronic News, 1998, three part series in issues of March 30, April 6 and April 13.

11. Semiconductor Industry Association, *National Technology Roadmap for Semiconductors*, 1997.

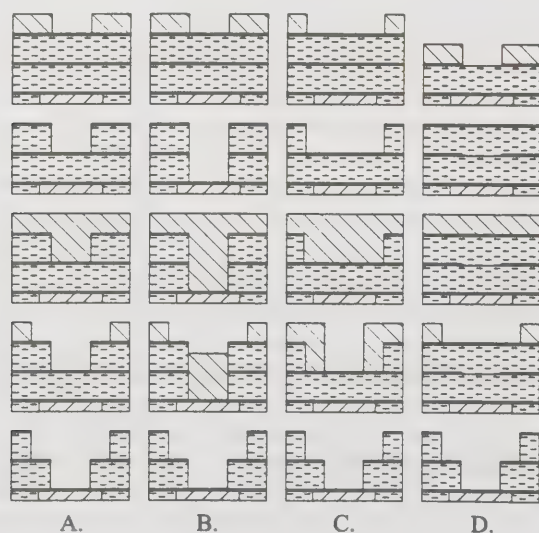


Figure 1. Dual damascene patterning strategies.

- A. Via first, partial etch,
- B. Via first, full etch,
- C. Trench first,
- D. Buried etch mask, from ref. [3].

 Photoresist
  Low k
 SiNx
  Copper

Advanced Packaging of Integrated Passive Devices for RF Applications

Elizabeth A. Logan, Howard M. Clearfield, James L. Young, Don H. Bolton

Intarsia Corporation
48611 Warm Springs Boulevard
Fremont, CA 94539
510-354-6301, 510-354-6330 (fax)
elogan@intarsiacorp.com

Abstract

Modern RF and wireless systems, in particular the portable and handheld consumer devices, are placing huge demands on component and assembly technologies. The integration of many different functions into one unit means that the component count has become very large. Since passive components form the majority of the devices mounted on the circuit board, there is considerable incentive to identify ways of reducing the number and/or size of these devices without compromising system performance. Interest is now focusing on the role that integrated passive devices (IPDs) can play in achieving some of these aims and a number of new approaches are emerging in this field. Integrated passive devices can provide a large number of passive components and functional blocks within a single device, reducing the overall complexity at the board level and improving the ease of assembly. In order to be effective, however, both the manufacturing and packaging technologies need to be selected to optimize performance, areal density and cost. Chip-scale packaging, combined with a thin film passive device technology and a large area format manufacturing capability offers a solution to these issues, as well as providing a package form factor compatible with existing surface mount assembly operation.

This paper describes the thin film integrated passive technology and chip-scale packaging techniques which are being used to produce small outline, high performance integrated passive devices for use in a range of RF, wireless and digital applications.

Introduction

The demands on today's wireless systems, both in terms of functionality and form factor, are driving the trend towards more compact, lightweight and better performance components and subsystems. In particular, this trend has prompted the conflicting requirements of a greatly increased component count together with a reduction in available printed circuit board area. While silicon integration has advanced rapidly in terms of providing increased functionality within a reduced device footprint, a similar trend has been lacking in the passive component arena. This is amply demonstrated by considering the circuitry within a cellphone, where the passive component count now outweighs the number of active integrated circuits by a factor of 20:1, with some products containing nearly one thousand passive components [1]. Most efforts in the passives arena have focused not on improving the functionality of the passive device, but on reducing the discrete component size, to a point at which 0603 (60mils x 30 mils, or ~ 1.5 mm x 0.75 mm) and even 0402 (40 mils x 20 mils, or ~ 1 mm x 0.5 mm) single components are now in use. Such components can improve the on-board component density but pose ever increasing problems for the assembly houses, due to their very small size. One alternative is to combine many of the

discrete components into one device – an integrated passive device or IPD – that can provide the same, if not better, performance while alleviating the assembly difficulties by providing a more easily manageable format. To this end, the methods used to package these IPDs play a major role in determining their acceptance in the marketplace. The devices must be presented in a format that conforms to the existing, generally surface mount, standards, while the packaging must add a minimum of overhead to the device in terms of cost and size, so as to maintain its competitiveness with the discrete component solutions. New packaging solutions are therefore needed to address these requirements.

This paper will discuss a new suite of technologies that are being employed to address the issues of size, weight, performance and cost. These technologies combine the use of integrated passive devices with an advanced packaging technique - chip-scale packaging - to give passive device networks and functional blocks which offer significant advantages over conventional discrete passive components and are compatible with current assembly methods.

Integrated Passive Devices

An integrated passive device is a device in which a number of passive components such as resistors, capacitors and inductors, are combined together on a single component. The concept of an integrated passive device is not new. A number of manufacturers have, for some time, offered small resistor (R) and resistor-capacitor (R-C) networks prepared using either thick film, such as cofired ceramic, or thin film processes. Traditionally, the thin film devices, which offer the most potential for dense integration and component precision, have been built using a silicon substrate to support the thin film process layers. Silicon provides a convenient format for thin film processing but, because of its semiconducting properties, begins to impose performance limitations at RF frequencies. It is also traditionally more costly to process and package than the thick film devices.

In silicon IC processing, some of the biggest cost savings to be made are those due to the economies of scale associated with the wafer size. This is the prime force behind the drive to increase wafer size from 6" to 8", and now to 12" diameter. A similar cost structure applies to thin film IPDs. For this reason, the device technology described here uses not silicon, but glass as the thin film device substrate. This choice provides two significant advantages. One, that glass is a low cost material which is available in large panel format, thus permitting very large scale and cost effective manufacture. Secondly, the glass chosen provides a high resistivity, low loss, low dielectric constant medium which is suited to higher frequency operation. Typical values for the material in use are $\epsilon_r = 5.7$, $\tan \delta \leq 0.001$.

The integrated passive devices are formed using semiconductor processing techniques to give precision thin film capacitors, resistors and inductors [2]. A schematic cross-section of the process structure is shown in Figure 1. Aluminum or copper/nickel/gold is used for the metal traces and Cyclotene® 4024 BCB [3] as the interlayer dielectric. The capacitor technology is based upon a thin Al_2O_3 dielectric, formed by the anodization of an aluminum layer. The capacitance density achieved by this technique is 500 pF/mm^2 with a tolerance of better than 5%. The range of values obtained with this process is generally limited to 3 - 500 pF. The lower limit is dominated by the ability to accurately pattern small areas using the photolithographic process. For smaller capacitor values, the BCB dielectric layer can be used to provide parallel plate capacitors. The upper limit is an arbitrary one. Much larger value capacitors can be built, but at the expense of substrate area.

Tantalum nitride is used as the resistor material. This gives films with sheet resistivities from 10 ohms per square to approximately 150 ohms per square with a temperature coefficient of resistance (TCR) between -50 and -150 ppm. In practice these films can be used to provide resistors with values from $\sim 1\Omega$ to a few tens of kilo-ohms. Spiral inductors from $\sim 1 \text{ nH}$ to $\sim 100 \text{ nH}$ with Qs of ~ 20 can also be patterned within the metal layers.

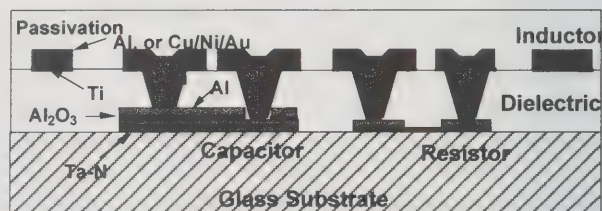


Figure 1: Cross section of the Al/BCB integrated passive device structure.

Given these capabilities, it is possible to form small circuits of passive components, including RF functional blocks, using these techniques. Such circuits have been demonstrated elsewhere using similar technologies [4,5] and their performance compares favorably with the standard discrete approach. One key to achieving the desired properties of the completed circuit lies in the successful generation of a cell library of RF characterized components, from which the individual parts of the circuit are selected. Figure 2 shows a small selection of the test components being used to generate the data for this technology. A full range of component test structures is currently in process, however preliminary



Figure 2: Test components used for process characterization.

results measured at frequencies up to 5 GHz show very promising results, in line with those described by other authors [4]. Another major factor in achieving the desired performance is the ability to ensure that components are identical across a panel or wafer. Properly done, thin film technology can assure this reproducibility. These benefits can be extended to the package by the adoption of a packaging strategy which utilizes the same technologies. Wafer- or panel-level chip-scale packaging can fulfill this role.

Chip-Scale Packaging

Chip-scale packaging is a term used to refer to any packaging scheme in which the package edge dimension does not exceed 1.2x that of the device inside [6]. These packages offer significantly higher wiring density than virtually any other packaging or assembly option, with the exception of flip chip, and take up far less area on the printed circuit board than conventional devices. They are capable of giving a very thin, low profile device, ideally suitable for portable applications. They also provide mechanical protection to the device within and can be mounted using a standard surface mount process, without the need for underfill, making them compatible with most assembly lines.

Many companies are currently working on chip-scale packaging and the number of process permutations used to achieve this package size is vast. Some, such as those produced by Fujitsu, rely on the addition of a leadframe to the device, others on the use of a substrate or interposer. This can be a rigid substrate, usually BT resin or FR4 such as those used by Sony or Motorola or can be a flexible polyimide interposer as used by Mitsubishi. A third class of chip-scale packaging is that which is carried out at the wafer level, performing processing steps on the wafer to form the finished, packaged devices prior to dicing into individual components. This approach has been developed by a number of companies, amongst them Tessera, Shellcase and ChipScale Inc.

For the devices described in this work the Micro SMT® package technology [6,7], developed by ChipScale, Inc., was used as a starting point. This particular technology utilizes thin film processing techniques to bring contacts from the front to the back of the device, so that the finished structure is only a small amount larger than the device itself. Such a technology eliminates the need for wirebonds to connect the device to the package leads and provides contacts which are extremely short in length. In an RF environment, such a structure can prove to be of great benefit in reducing the series inductance associated with the package.

Figure 3 shows the process steps involved in fabricating a chip-scale packaged IPD. Although the process is performed on the whole wafer or panel, the diagram shows only one device. After completion of the steps shown, the whole substrate is diced into the individual components.

The first stage in the process is to (a) apply a uniform coating of polymer to the entire substrate. This helps to provide both mechanical integrity during processing and environmental protection to the completed device. In cases where thermal coefficient of expansion (TCE) mismatches are large between the substrate and the polymer, a cap of the same material as the substrate can be added above the polymer to provide a symmetrical structure. The substrate is then thinned to approximately 100 microns thick (b), and a tapered saw used to provide trenches in the glass (c) exposing the

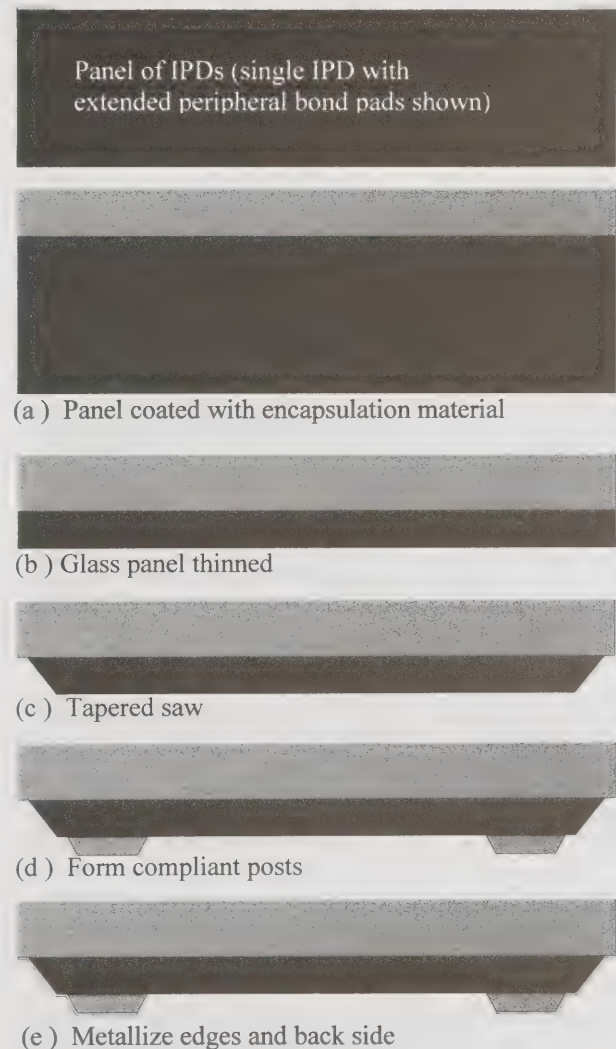


Figure 3: Process steps in the formation of a chip-scale packaged integrated passive device. The diagrams show just one device out of a panel.

lead metallization. The tapered profile is important in that it allows good step coverage during deposition of the subsequent metallization so that electrical continuity is achieved from front to back of the device. Figure 4 shows a micrograph of the back surface of a glass substrate at this stage of the process. For larger devices, the next step is to add a compliant epoxy bump to the location of each pad on the underside of the device (d). For very small devices, where the CTE mismatch between device and board is less likely to impact the reliability of the assembled part, these bumps can be omitted. The backside of the device is then metallized and patterned to create the "wraparound" leads (e). As these leads are plated with copper, nickel and gold their electrical resistance is comparable to a printed circuit board trace. This technique has the advantage of giving a final, face-up format for the package, with leads that extend up the sides of the device. The finished device therefore closely resembles a peripheral leaded surface mount component and, after assembly to a board, shows a solder fillet which can be inspected by conventional techniques. For higher pin count devices, where the pitch of the peripheral leads becomes too fine for standard assembly equipment, it is possible to re-route the pads on the bottom of the device to give an area array footprint, resembling a ball grid array (BGA). Figure 5 shows a number of chip-scale integrated passive devices compared to traditional packaged devices, from which it is possible to appreciate the area efficiency of this approach.

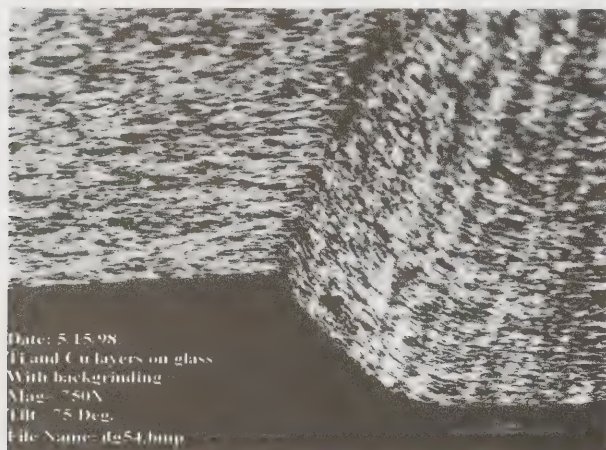


Figure 4: Trench sawn in glass to facilitate the backside to front-side contacts for the chip scale process.

Conclusions

It is apparent that there is a significant role that IPDs can play in achieving some of the more demanding requirements of wireless systems. The approach can offer not only a simplified assembly scheme and board layout for the end product, but also the potential for improved performance through the control of

component interconnections. Despite the potential advantages, however, it has to be recognized that the majority of applications for such devices fall within the consumer market, where cost is a prime driver in selecting the technologies and materials. For this reason it is imperative that the manufacturing and packaging methods used in producing such devices be selected so as to keep the cost of the IPD in line with that of its discrete counterparts. The combination of a panel scale process with a wafer- or panel-level chip-scale packaging strategy, such as that described above, offer the potential for considerable cost reductions, making the thin film IPD a viable approach in many modern systems.

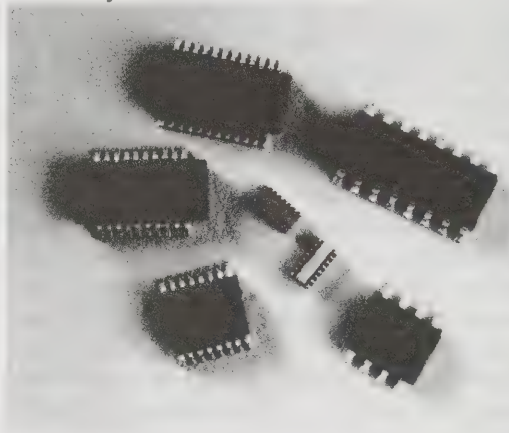


Figure 5: Chip scale integrated passive devices (devices courtesy of ChipScale, Inc.) compared to standard SMT components.

References

- [1] R. Redolphy and M. Christensen, "Inside a state-of-the-art mini cellphone", *Portable Design*, Jan. 1997, pp18-24
- [2] H. M. Clearfield et al., "Integrated Passive Devices using Al/BCB Thin Films", *Proceedings International Conference and Exhibition on MCM & High Density Interconnect*, Denver, CO April 1998 pp 523-527
- [3] Cyclotene® is a registered trademark of The Dow Chemical company, Midland, MI
- [4] D. Benson et al., "Integrated Passive Components for RF Applications", *Proceedings, 1997 Wireless Communications Conference*, Boulder, CO pp 175-180
- [5] R. G. Arnold, C. C. Faulkner and D. J. Pedder, "MCM-D and Direct Module Attach for RF Applications", *Proceedings International Conference and Exhibition on MCM & High Density Interconnect*, Denver, CO April 1998 pp 286-290
- [6] T. W. Goodman and E. J. Vardaman, "CSP Markets and Applications", *TechSearch International Inc.*, publ. Jan 1998
- [7] Micro SMT® is a registered trademark of ChipScale Inc., San Jose, CA.

Quick Prototyping of Flip Chip Assembly with MEMS

Ronda Irwin, Wenge Zhang, Kevin Harsh, Y.C. Lee

University of Colorado at Boulder

Campus Box 427

Boulder, CO 80309

(303)-492-3393, (303)-492-3498 (FAX), leeyc@colorado.edu

Abstract

This paper describes a process to transfer Microelectromechanical Systems (MEMS) devices to a secondary substrate using flip-chip thermosonic bonding. A standard wire-bonding machine was used to place ~100- μm bumps on unreleased MEMS chiplets. The bumped chiplet was then flip-chip bonded to a secondary substrate containing a microwave coplanar wave-guide (CPW). After bonding, the entire assembly was run through the MEMS release process, after which the MEMS host substrate was removed. The thermosonic bonding was a very reliable prototyping tool with a 100% bonding yield. The transfer process can be used with any MEMS that can be wire bonded. The process can also be applied to a variety of applications.

Key words: MEMS, Flip-chip, thermosonic bonding, transfer bonding

Introduction:

Microelectromechanical Systems (MEMS) are becoming critical components for sensors, optoelectronic devices, microwave circuits, and a variety of other high technology systems. Combining MEMS with existing technologies, however, can be difficult. Standard processes, such as some of those used to build Multi-chip Modules (MCM), may not be compatible MEMS. Flip-chip assembly, however, is a hybrid technology that can integrate MEMS into other technologies to meet new specifications. This study demonstrated how to use flip-chip to fabricate a MEMS tunable capacitor for millimeter-wave applications.

Recent studies at the University of Colorado have found that MEMS used as capacitors do not behave as expected. Figure 1 is a diagram of an on-chip capacitor measurement. The MEMS actuator was the top plate in a parallel-plate capacitor. The bottom plate was fixed to the substrate directly below the top plate. Figure 2 presents the measured capacitance at 13 MHz with different DC biased voltages applied to the MEMS actuator. With the applied voltage, we expected to have a slight increase in the capacitance because of the gap reduction by electrostatic force between the capacitor plates. Surprisingly, the capacitance decreased significantly upon application of DC bias. After investigation, it was understood that there was strong coupling among the measured capacitance, C_m , and other spurious capacitive loadings associated with the MEMS device, as shown in Figure 1. As a result, we decided to completely separate the MEMS device from its

host substrate to eliminate unwanted capacitances, as illustrated in Figure 3.

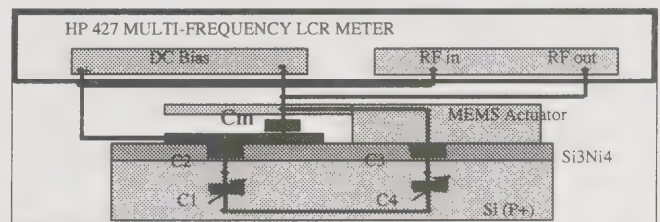


Figure 1. MEMS actuator as part of a parallel-plate capacitor that is connected other capacitors through the silicon.

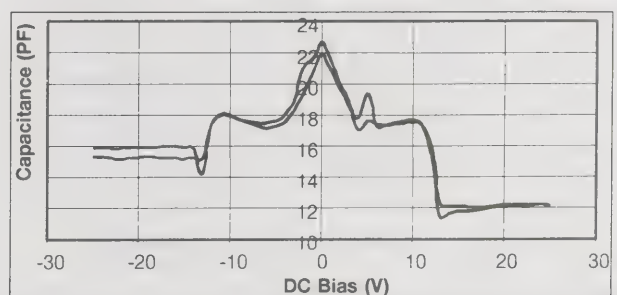


Figure 2. Measured capacitance for a MEMS parallel-plate capacitor.

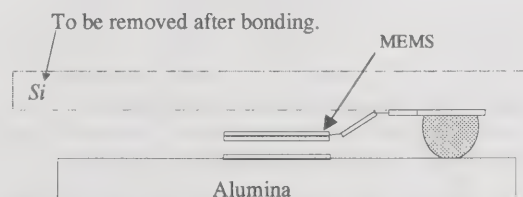


Figure 3. MEMS with silicon substrate to be removed.

To separate the MEMS device from its host substrate, we used a flip-chip bonding process. The process was similar to that of Singh, et al., which used flip-chip compression bonding to transfer released MEMS devices to a secondary substrate [1]. The process described in this paper, however, transferred unreleased MEMS to a secondary substrate using flip-chip thermosonic bonding. Bonding unreleased MEMS minimized the potential to damage the devices during bonding. Thermosonic bonding was compatible with wire bonding technology and could be easily adapted in most packaging laboratories. After the bonding, a release process was developed to remove the host silicon substrate without damaging the transferred MEMS device on the secondary substrate.

Transfer Process

Figure 4 illustrates the transfer process that is described as follows

1. Placed $\sim 100\text{-}\mu\text{m}$ gold bumps on all the transfer pads and one anchor pad.
2. Pressed gold bumps against smooth glass for bump planarization.
3. Thermosonically bonded MEMS to a secondary substrate.
4. Removed SiO_2 in HF.
5. Removed MEMS host silicon substrate.

The transfer pads were only connected to the host substrate with the sacrificial SiO_2 layer. When the SiO_2 was dissolved in HF, the transfer pads, and thus the MEMS device, was completely disconnected from the host substrate. The anchor pad was

permanently connected to the MEMS host substrate. The bump on the anchor pad held the host substrate in place during processing to protect the transferred MEMS device. Figure 5 shows the MEMS chiplet used study and where the transfer and anchor bumps were placed. The MEMS consists of a polysilicon plate $500 \times 500\text{-}\mu\text{m}$ and $1.9\text{-}\mu\text{m}$ thick that is connected to the transfer pads with thin polysilicon wires. The thin wires also act as thermal actuators to deflect the plate vertically, as described in [2].

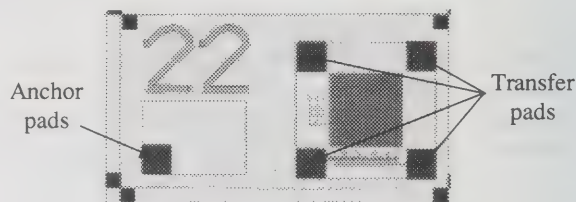


Figure 5. MEMS device transferred.

The secondary substrate was a microwave coplanar wave guide (Figure 6). A $50\text{-}\Omega$ transmission line was patterned and etched into $4\text{-}\mu\text{m}$ thick gold on an alumina substrate. The pattern also included transfer-bonding pads and capacitor plates.

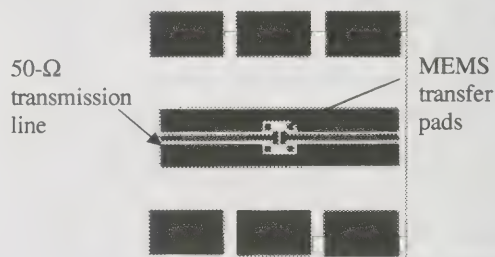


Figure 6. CPW design on secondary substrate.

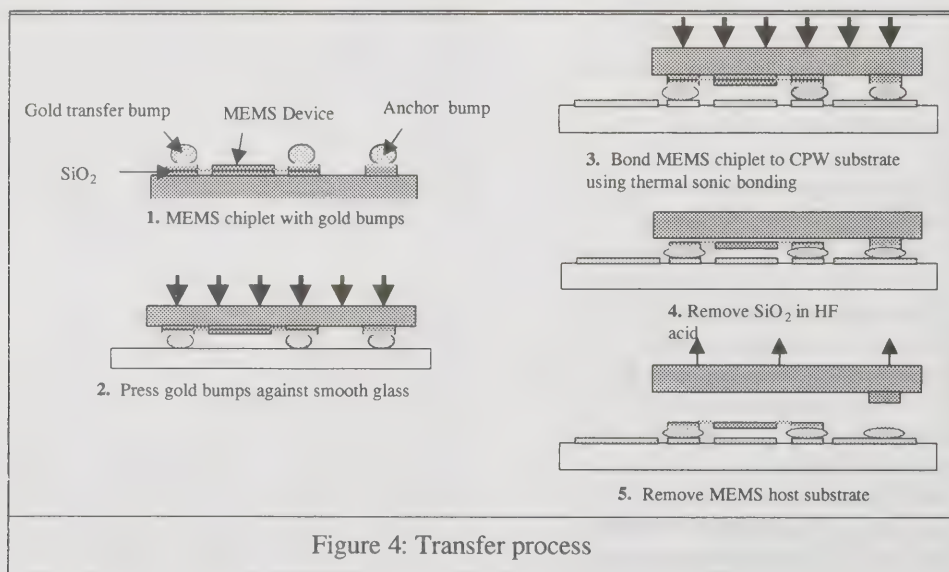


Figure 4: Transfer process

Before bonding the MEMS chiplet to the CPW, the gold bumps on the MEMS chiplet were pressed against a smooth glass substrate using 4-lbs of force. The pressing step was required to flatten any trailing wires from the top of the bump left by the wire-bonding machine. In addition, pressing the bumps provided better planarity across the bumps during the actual bonding. After the bumps were pressed, the MEMS chiplet and the CPW substrate were aligned with respect to each other in the thermosonic-bonding machine. The thermosonic flip-chip bonding technology is described by McLaren et al. [3]. The MEMS chiplet was then bonded to the CPW substrate using the following parameters:

1. 160°C bonding stage temperature.
2. 4.5-lbs bonding force.
3. 4.5 ultrasonic for 500-ms.

These bonding parameters yielded an approximate 20- μ m gap between the transferred MEMS and the CPW substrate.

For the release process, we dissolved the sacrificial SiO_2 layers in 48% aqueous hydrofluoric acid (HF) for 4 minutes. The MEMS device was separated from its host substrate after release. The 4-minute release time was longer than the 2.5-minute suggested by MCNC. More time was needed to allow the HF to diffuse under the 20- μ m gap between the MEMS chiplet and the CPW substrate. In addition, there were no etching holes in the MEMS transfer bonding pads to facilitate etching under the gold bumps. After the HF release, the entire assembly was processed according to the following steps to minimize stiction.

1. Soak in 4:1 methanol to water for 10 minutes.
2. Soak in methanol for 20 minutes.
3. Dry in a 110°C oven for 15 minutes.

Neither the gold CPW structure nor the alumina substrate suffered observable degradation during the release process. After the release process, the MEMS host substrate was removed by breaking the anchor with clamps. Figure 7 is a scanning electron micrograph (SEM) of a successfully transferred MEMS device on a CPW substrate. The MEMS plate is connected to the two bonding pads on the right. The two bonding pads on the left provided a symmetry during bonding and are not connected to the plate. The capacitor was formed between the MEMS plate and the two metal pads on the substrate.

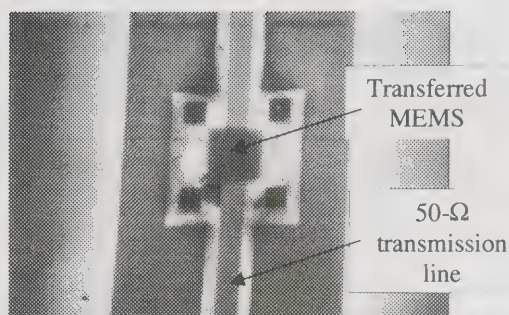


Figure 7. Successfully transferred MEMS device on a CPW substrate.

Results and Discussion

The thermosonic bonding process was very reliable, as all of the chiplets tested in this study were successfully bonded. The bonding is an excellent prototyping tool for MEMS transfer. Some transferred devices, however, were misaligned with respect to the CPW. The misalignment occurred when the MEMS chiplet slipped on the CPW substrate during the ultrasonic blast. Figure 8 is an SEM that shows a misaligned MEMS device. The problem could be solved by rigidly holding the MEMS chiplet during the thermosonic bonding. An alternative solution was to use solder reflow bonding instead of thermosonic bonding. The self-aligning molten solder would pull the MEMS device into near perfect alignment with respect to the CPW.

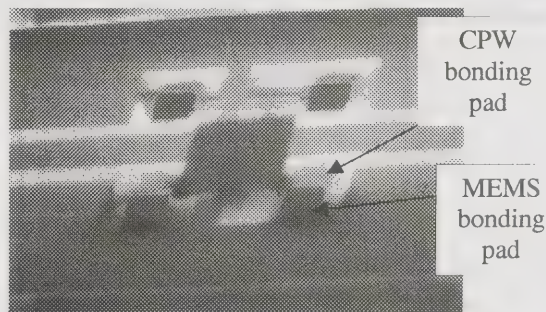


Figure 8. SEM of a misaligned MEMS device.

In addition to the misalignment problem, there was another problem observed during the release process. Some of the transferred MEMS devices were washed away in the release solutions. As a result, care was taken during the release process to slowly move the assembly from one bath to the next. The slow motion prevented the onrush of fluid that washed away the transferred device. It should be noted that the wash-away problem was more severe when the MEMS host substrate was not fixed in place during release. Without fixing the host substrate,

ninety percent of transferred MEMS devices were lost in the release solutions on test structures. The use of the anchor pad was essential to assure a successful substrate removal.

Further tests of the release process led to the discovery of another phenomena that can greatly effect bonded MEMS. Due to the transfer process, the release of the MEMS structures must be done with the host substrate attached. This limits the flow of the HF solution, which in turn increases the total required etch time. As tests were done, it became apparent that the resistivity of the MEMS structures was increasing considerably as the etching progressed. A simple test was done using a simple MEMS cantilever beam fixed at both ends. The beam was subjected to the release process and the resistance was measured periodically. The beam was large enough to preclude the possibility that the increase was an affect of the cross sectional area being increase. The results of three tests are shown in figure 9. This increase is significant enough to be a possible problem for some MEMS applications. Two possibilities for why this is happening are presented here. First, the MUMPS process used to fabricate the MEMS used here uses phosphorus doped polysilicon as the structural layer. The phosphorus is mainly contained in a thin outer layer of the polysilicon. This doped area contributes a majority of the conductivity of the polysilicon. By etching the MEMS for a time much longer than that prescribed by MCNC, we are eliminating a significant amount of this conductive area. The second possibility is that bonding the MEMS to an associated circuit is causing the etch rate to be sped up electrochemically. This is a commonly observed phenomenon in MEMS. It often causes the silicon opposite to the metalized areas to be significantly over-etched. At this point, it is not certain if either of these possibilities, or a combination of both, is the cause of the resistivity problem. A solution may be to use a partial pre-etching step before the MEMS are bonded. This would have the effect of cutting down on the amount of time electrochemical etching can occur, and reduce the overall etch time.

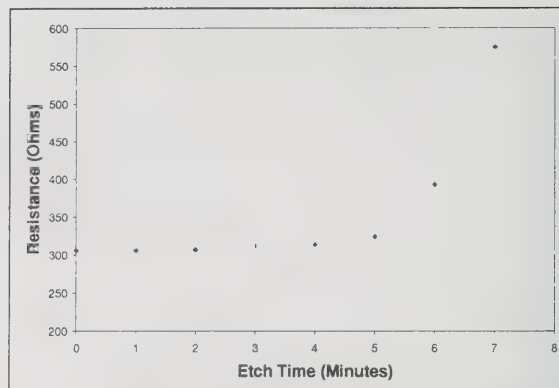


Figure 9: Increase of resistance with respect to etch time for a beam fixed at both ends.

Conclusions

Successful transfer of MEMS devices to a microwave CPW has been achieved. The MEMS were bonded before release, which resulted in no observable bonding damage to the MEMS devices. Transferred MEMS showed excellent performance in the microwave circuits. The thermosonic flip-chip bonding used for the transfer is well suited for quick prototyping of MEMS devices in microwave and millimeter wave applications. It is compatible with all the MEMS that are wire bondable, and can also be used in a variety of MEMS transfer applications.

Acknowledgements

This project was supported by the Department of Defense (MDA904-97-C-0320).

References

1. A. Singh, D.S. Horsley, M.B.Cohn, A.P. Pisano, R.T. Howe, Batch Transfer of Microstructures using Flip-Chip Solder Bump Bonding," Transducers 97: 1997 International Conference on Solid-State Sensors and Actuators, Chicago, June 16-19, 1997.
2. Comtois, J.H. and Bright, V.M., "Surface micromachined polysilicon thermal actuator arrays and applications", Proc. 1996 Solid State Sensor and Actuator Workshop, Hilton Head, SC, pp. 174-177, 2-6 June, 1996.
3. McLaren, T., Kang, S.Y., Zhang, W., Ju, T.H., Lee, Y.C., 1997, "Development of a Thermosonic Bonding Process for an 8 x 8 VCSEL Array Based Optical Transceiver", IEEE Transactions on Components, Parts and Manufacturing Technology Part B: Advanced Packaging, vol. 20, no. 2, May

Partially Interdigitated Combline Filter

Reddy Vangala
Motorola CPG
Ph: (505) 348-4387

Abstract

Ceramic combline filters have been the preferred choice for duplexing the transmitter and the receiver to the antenna in wireless portable applications. However, it is also the biggest component in today's radios. Therefore, there is great interest in reducing the size of these filters. Low cost is also a big driver. One incremental solution is to find a way to eliminate the external metal shield that, typically, accounts for 10% to 20% the filter volume and for 5% to 10% of the manufacturing cost. However, the external metal shield helps to reduce the unwanted parasitic field coupling between combline resonators at the open circuit end. These fields terminate on the ground plane provided by the shield. The solution presented here is to interdigitate some of the resonators in the filter and thereby distribute the ground plane sufficiently to eliminate the shield without adversely affecting the stopband attenuation performance.

Introduction

Due to the crowded radio frequency spectrum, the duplex filters used in today's wireless phones are required to be very frequency selective. They reject a wide range of undesirable signals from interfering with the receiver and to contain the spurious radio emissions from the transmitter to protect the receiver and also to comply with regulations. They must have low insertion loss to enable long talk-time and good receiver sensitivity. Small size is desirable and much needed in the new multi-band phones, where the component count is higher. Ceramic filters have been satisfying these needs for many years in cellular phones operating at 1 GHz and more recently in PCS phones operating at 2 GHz.

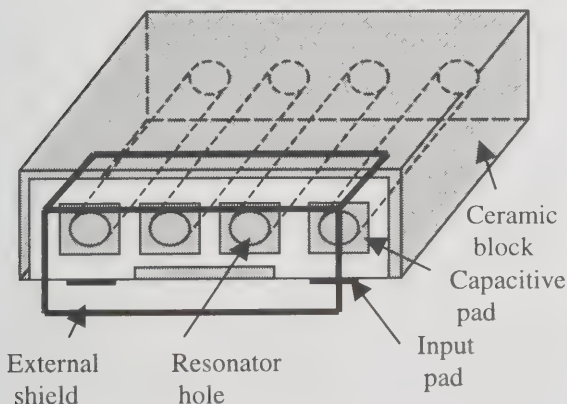


Fig.1. Ceramic combline filter with shield

Ceramic combline filters, as shown in Fig.1, consist of TEM mode transmission lines, implemented in a high dielectric constant material. All of the lines are short-circuited at one end of the ceramic body and open-circuited at the opposite end. Furthermore, the lines must be less than one quarter wave long at the frequency of operation. The remainder of the 90° electrical length required for the resonance is satisfied by loading capacitors at the open end of the lines. The resonance condition is decided by

$$Z \tan(\theta) = 1/\omega C$$

Where Z = characteristic impedance of the line

$$\omega = 2\pi f = \text{radian frequency}$$

$$\theta = \text{electrical length of the line at } f$$

$$C = \text{loading capacitance}$$

The lines are realized as metallized holes in ceramic and the loading capacitance is formed as a gap between printed pad, attached to the open end of the line, and the ground electrode. The ground electrode is the conductor covering most of the ceramic block with the exception of the open circuited end.

At the open-circuited end, parasitic coupling between non-adjacent resonators leads to poor stopband attenuation performance. In all pole combline filters, it is possible to provide ground bars between resonators to keep this parasitic coupling to a minimum. However, most of the high selectivity filters, where the stopbands are very close to the passband edges, need finite transmission zeros to improve the stopband rejection. Such zeros are easily implemented with the addition of a capacitor between adjacent resonators. The capacitor will parallel resonate with the inductor realized within the block, thus

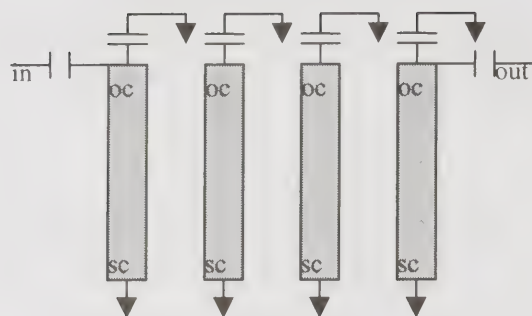


Fig.2a. Combline topology

presenting an open circuit between the two adjacent resonators at the prescribed frequency. Unfortunately, the presence of this capacitor between adjacent resonators precludes the use of a ground bar. Designers have resorted to the external metal shield, which essentially encloses the open end of the block thus providing ground plane for the parasitic capacitors to terminate on. The shield has to be close enough to the open surface of the block in order to be effective and yet far enough not to disturb the loading capacitors during the manufacturing process. For these reasons it is, typically, set about 1mm off the block. This can account for as much as 20% of the board area occupied by the filter, for filters designed to operate at high frequencies. Additionally, it also adds to the filter height above the board and contributes to the yield losses in manufacturing. For these reasons, it is of interest to find alternate ways to minimize the parasitic non-adjacent capacitive couplings.

Operation

Fig. 2a , shows the schematic of a typical combline filter. All of the transmission lines are short- circuited at the bottom and open-circuited at the top. In ceramic block implementation, this topology leads to a lot of parasitic coupling. The interdigital filter schematic, shown in Fig.2b, uses quarter wavelength resonators with alternating open and short circuits. The lines can be 90° long making the loading capacitors unnecessary. This topology helps to distribute the ground plane evenly on both ends of the ceramic block. Therefore, it is possible to realize good ultimate stopband attenuation without using an external shield. However, the interdigital filter doesn't support the transmission zeros between adjacent resonators and hence the selectivity is not as good as the combline filters.

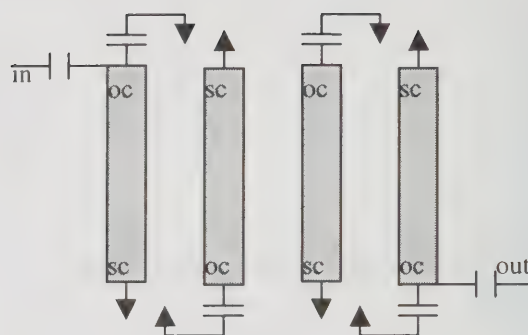


Fig.2b. Interdigital topology

What is proposed here is a hybrid between the combline and interdigital topologies. Fig. 3, shows a 4 pole filter that is essentially a combline filter, where some resonators are disposed in interdigital fashion. Thus, some transmission zeros that are critical to the steep selectivity are preserved and at the same time the internal ground plane is distributed more evenly between both ends of the ceramic body as in the interdigital filter. This reduces the need for external metal shield. The arrangement allows ground plane between non-adjacent resonators while keeping some of the adjacent resonator transmission zeros. In a pure

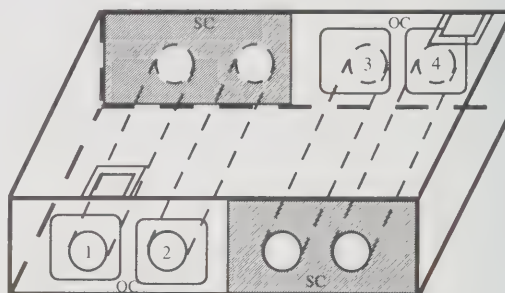


Fig.3. Four pole partially interdigitated combline

comline filter the number of such transmission zeros is limited to $N-1$, while in the proposed structure it is limited to $N/2$ and $(N-1)/2$ for even and odd number of resonators, respectively.

Typically, coupling between interdigital resonators is quite strong compared to the combline resonators. That means for a given bandwidth, the spacing between interdigital resonators has to be bigger resulting in a bigger filter structure. There are several ways to solve this problem. One way is to shorten the resonators sufficiently but this will degrade the resonator unloaded Q . Another way

might be to offset the resonators in the resonator length direction. However, that would result in non-uniform ends of the block and hence difficult to manufacture. In Fig. 3, the inter-resonator coupling is reduced by placing a pair of ground holes (not shown in the figure) midway between the two interdigital resonators. Two smaller ground holes, instead of one bigger ground hole, seemed to improve the resonator unloaded Q. The loading and I/O coupling capacitors are realized using fired thick film conductor gaps.

Fig. 4 shows the picture of a PCS prototype filter built using the concept. It is a 4 pole filter with two transmission line zeros located below the passband. Resonator pairs 1,2 and 3,4 are configured in combline fashion while the pair 2,3 is disposed as an interdigital filter. It is built using a ceramic with a dielectric constant of 37. The passband is centered at 1880 MHz and designed to have a 60 MHz bandwidth with an insertion loss less than 3 dB. More than 50 dB attenuation is required at 100 MHz below the center frequency. The filter is modeled with Touchstone circuit simulator. All of the transmission line impedances are calculated using LINPAR, a commercially available 2D electrostatic solver. The capacitor values are optimized in the circuit simulator and both of the transmission zeros are placed at 1780 MHz for the best attenuation response. The transmission zero above the passband is the result of parasitic coupling between the two IO pads. It is modeled as a small capacitor connected between the input and output terminals.

Results

Simulated and measured responses are shown in Fig.5. The passband insertion loss is 3 dB and the rejection below 1780 MHz is better than 60 dB. In a conventional combline filter, without external shield, it would not have been possible to get more than 40 dB of rejection. Simulated and measured responses agree fairly well at the passband and up to a few hundred MHz on either side of the passband. At the low end the measured rejection is typically better than the model but at the high end the measured rejection is worse than the model, due to the excitation of waveguide modes in the ceramic. The model assumes pure TEM propagation in the filter. Size of this filter is 0.44"x .213"x .15". The conventional combline filter with shield would occupy roughly .44"x .250"x.160". In higher dielectric constant implementations, the size difference could be even more significant.

Summary

A new hybrid filter topology is presented that can eliminate the need for an external metal shield in high performance filters, potentially reducing the volume and cost of the filter. The concept is illustrated with a simple 4 pole filter but can be easily extended to more complex multi-pole duplexers used in wireless communications.

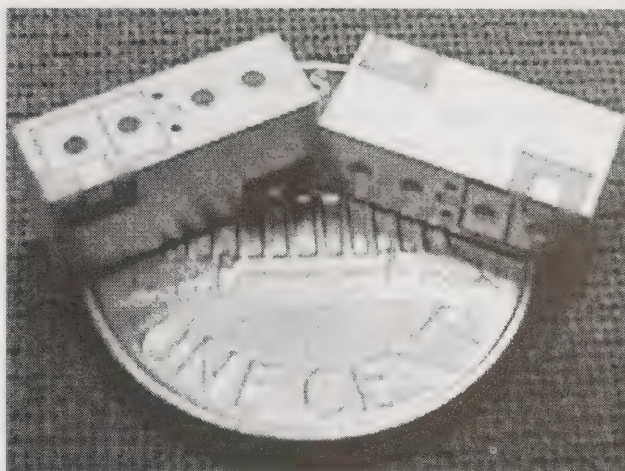


Fig.4. A four pole PCS prototype filter

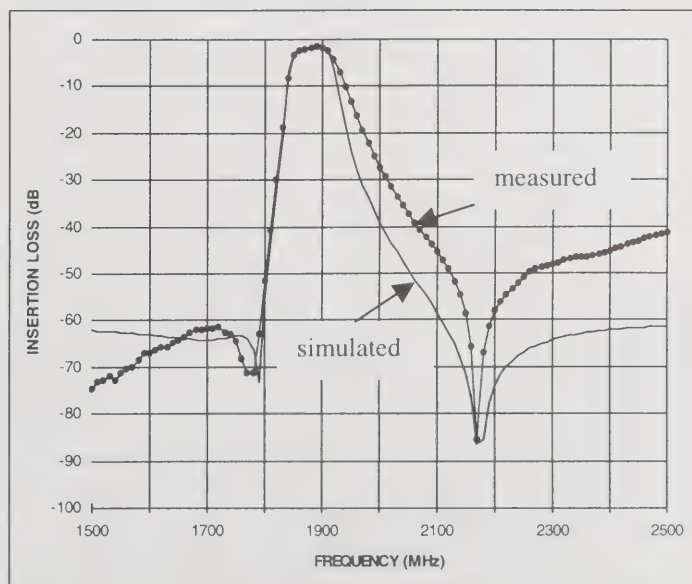


Fig.5 Simulated and measured responses of the filter in fig. 4.

An Improved Lumped-Element Equivalent Circuit for on Silicon Integrated Inductors

Paolo Arcioni, Rinaldo Castello, Luca Perregrini, Enrico Sacchi
University of Pavia, Department of Electronics – via Ferrata, 1 – 27100 Pavia, Italy
+39 382 505200, +39 382 422583 (fax), e.sacchi@ele.unipv.it

Francesco Svelto
University of Bergamo, Department of Engineering – viale Marconi, 5 – Dalmine (BG), Italy
+39 35 277308, fsvelto@ipvsm6.unipv.it

Abstract

This paper presents an improved lumped-element equivalent circuit of silicon integrated inductors, that accurately describes the effect on the Q-factor of the electromagnetic coupling between the metal strips and the substrate. In the proposed model a magnetic coupling describes the interaction between the substrate and the spirals of metal. The equivalent circuit is extracted using a characterization procedure that, from two-port wide-band measurement of the S-parameters, automatically yields the value of the lumped-element via the method of Least Square Minima.

Key words: radio-frequency, integrated inductors, equivalent circuits, Q-factor.

Introduction

The continuous evolution of the silicon technology is introducing some innovations in the RF design of transceivers for wireless communications, in particular with the introduction of on-chip spiral inductors. Nowadays technology scaling allows CMOS and BiCMOS processes to operate at radio-frequency and a great effort is under way to obtain a monolithic solution which meets specs set by mobile telecommunication standards (GSM, DCS-1800, DECT...). The goal of a CMOS or BiCMOS single chip, with low power consumption and reduced costs of fabrication, can be met only if these technologies will succeed in replacing Bipolar and GaAs in the RF analog part. Therefore, in a single chip, also inductors need to be realized on Si substrate along with all the other devices.

Integrated inductors on GaAs have been successfully realized, with Q-values of the order of tens. The lower substrate resistivity of Si, especially for CMOS processes, reduces the Q-values, due to the increased substrate losses. For these reasons it is important, in CAD applications, to model these effects by an accurate lumped-element equivalent circuit, deduced from experimental or numerical characterization.

This paper discusses this topic, and presents an improved lumped-element equivalent circuit for Si integrated inductors. Experimental data are also presented.

Integrated Inductors

The layout of a simple integrated inductor is shown

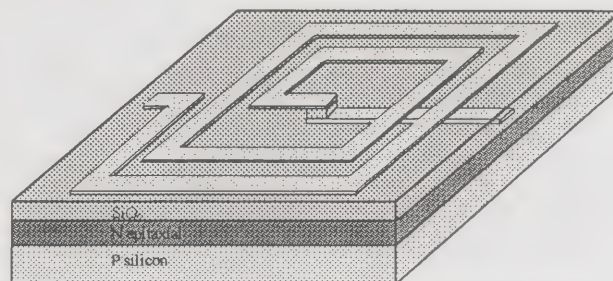


Fig. 1: layout of an integrated spiral inductor.

in Fig. 1: the top level of metal (the less resistive) is used to realize the spiral and a second level of metal is used to give access to the second port of the device. The substrate is silicon, with a thin epitaxial layer.

As these devices should be used in low-cost consumer applications, the technology used should be completely standard. So in their design we must distinguish between design parameters and process parameters. The former (area, shape, width of metal strip, distance between metals and number of turns, ...) can be manipulated by the designer to optimize inductors features; the latter (substrate resistivity, metal line conductivity, ...) are set by the technology. This constraint is very critical because silicon substrate is intrinsically characterized by high losses; CMOS substrate in particular, compared to the BiCMOS and to the Bipolar, is characterized by a very low resistivity. Metal resistivity, moreover, contributes to a further degradation of the quality factor of these devices.

Simple spirals inductors realized in standard technology are characterized by a quality factor, defined as [1]:

$$Q \equiv 2 \cdot \pi \cdot \frac{\text{energy stored}}{\text{energy loss in one oscillation cycle}} \quad (1)$$

Q values near 5 were reported [2] with inductors realized in a CMOS standard process. These values are too low to use these devices in commercial applications and, actually, an effort is needed to increase the quality factor.

To improve the Q-factor of the integrated inductors it is important to study their physical behavior. This would allow the understanding of the effects that limit their performances, and, consequently, it will be possible to optimize their design. Actually, it was not demonstrated whether the Q-factor is limited mainly by substrate or metal losses. In fact some authors [3] think that metal series resistance causes a very important limitation in the inductors features, many others [2], [4] affirm that the low resistivity of the CMOS substrate is the most serious problem.

For this reason, an equivalent circuit which allows to evaluate separately these two dissipative terms, should be very important in the optimization of these integrated devices.

The Improved Equivalent Circuit

Since the geometrical dimensions of these devices are small compared to the wavelength, it is possible to model them by a lumped-element circuit.

The equivalent circuit usually found in the literature is illustrated in Fig. 2. The lumped-element which constitute this circuit respectively represent:

- R_S : the series resistance caused by the finite conductivity of the metals;
- C_F : the fringing field capacitance between each side of the spiral;
- C_{OXin} , C_{OXout} : the capacitive effect of the oxide layer;
- C_{SUBin} , C_{SUBout} : the capacitive coupling with the substrate;
- R_{SUBin} , R_{SUBout} : the losses associate to the electrical coupling with the substrate.

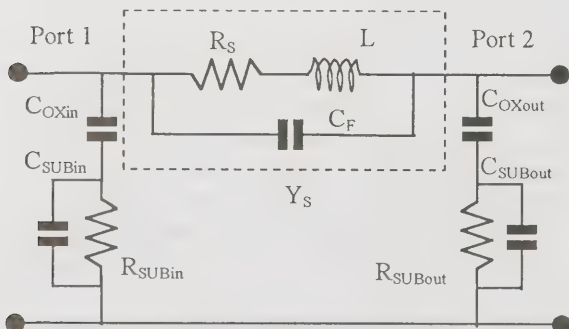


Fig. 2: the classical equivalent circuit.

This model seems not to be adequate in the accurate description of the coupling with the substrate, because it neglects completely the magnetic coupling between the spiral and the substrate. In fact the electromagnetic field induced by the current on the spiral induces opposite currents in the low resistive CMOS substrate. So, to better describe this coupling, we inserted in the equivalent circuit a substrate resistance (R_{SUB}), coupled through a transformer to the inductor (Fig.3). It is important to observe that, though apparently we introduced two new lumped parameters (N_C and R_{SUB}), the elements are equivalent to a shunt resistance equal to $(N_C)^2 \cdot R_{SUB}$, and the transformer is introduced only to better describe the physical phenomenon.

This effect, which probably could be neglected in the GaAs or Bipolar inductors, is very important in CMOS technology; in fact CMOS substrate is characterized by a so high doping concentration that in the most modern technologies the equivalent resistivity of the substrate is close to $0.1 \Omega\text{cm}$. This causes a loss angle which reaches values close to 1 Rad in the frequency range of interest.

Characterization and Validation of the Circuit

We designed and realized some inductors both in BiCMOS and CMOS technology. Their geometrical characteristics are reported in Table 1, where the column on the right reports the predicted inductance value, calculated on the basis of [5]. In [6] we characterized these inductors on the basis of the equivalent circuit of Fig. 2. The values are obtained using the de-embedding procedure described in [6], based on a two-port wide-band measurement of the S-parameters of the integrated inductors and of a reference structure (see Fig. 4). The peculiarity of this de-embedding procedure, based on the π -topology of the circuit, is the iterative extraction of simple circuital blocks, which makes computationally easier and more accurate the extraction of the lumped-element values. Because also the new equivalent circuit has a π -topology, we can repeat the same procedure, extracting the lumped-element of this new circuit.

The parameter's extraction is accomplished by means of

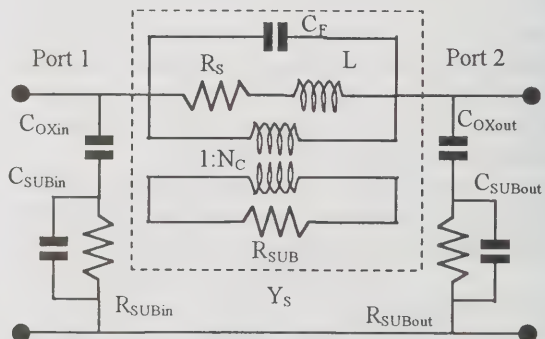


Fig. 3: the improved equivalent circuit.

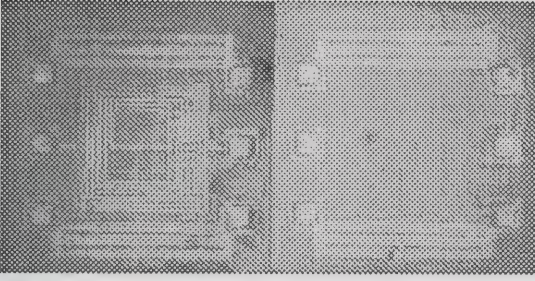


Fig. 4: top view of an integrated inductor and of a the test pattern.

a numerical code based on the Least Square Method and the values of the admittances are derived minimizing the difference between the magnitude of measured and modeled admittances. In particular we focus our attention on the Y_S admittance (see Fig. 3).

Fig. 5 shows the good fitting between the new model and the measurements in the whole frequency range of interest. On the contrary the old equivalent circuit fails around the self-resonance frequency, the frequency at which the magnitude of Y_S has a minimum. Physically it corresponds to the resonance of the inductance L with the fringing field capacitance C_F and it sets the maximum frequency at which these devices can be used.

We then extract all the unknown lumped parameters (values relative to the Y_S block are reported in Table 2) and we obtain the whole characterization of the inductors.

Quality Factor

It is interesting to analyze the quality factor of these devices. After a two-port characterization of the inductors we can deduce this parameter from the equivalent circuit of Fig. 3. The quality factor is computationally defined as the ratio of the imaginary part to the real part of the input impedance at port 1 when port 2 is short-circuited [2]. We have:

$$Q = \frac{\text{Im}\left(\frac{1}{Y_{11}}\right)}{\text{Re}\left(\frac{1}{Y_{11}}\right)} \cong \frac{\omega \cdot L \cdot R_{\text{SUB}}}{R_S^2 + R_{\text{SUB}} \cdot R_S \cdot \omega^2 \cdot L^2} \quad (2)$$

	Length	Width	Space	N	L
CMOS					
Ind1_c	300 μm	9 μm	4 μm	2	2.99nH
Ind2_c	300 μm	9 μm	4 μm	4	7.83nH
Ind3_c	300 μm	9 μm	4 μm	6	12.3nH
BiCMOS					
Ind1_b	300 μm	24 μm	4 μm	2	1.83nH
Ind2_b	300 μm	14 μm	4 μm	2	2.51nH
Ind3_b	300 μm	14 μm	4 μm	4	5.96nH
Ind4_b	300 μm	9 μm	4 μm	6	12.3nH

Table 1: inductor's geometrical characteristics.

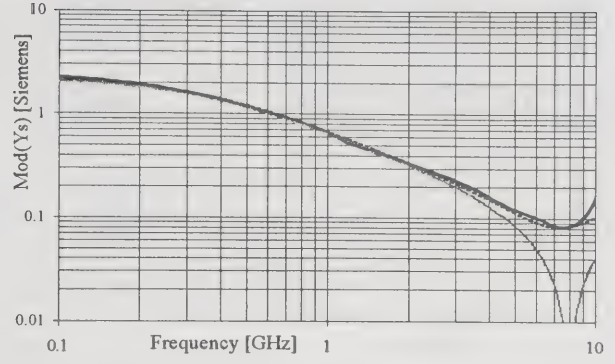


Fig. 5: comparison between the measured $|Y_S|$ vs. frequency (continuous line) and the deduced from the old model (gray) and from the new model (dotted line).

The Q-values obtained for the inductors we measured are rather low. In fact, as reported in Fig. 6, is observed that the maximum Q deduced from measurements is close to 4 for BiCMOS inductors and close to 3 for the CMOS ones. Nevertheless, from our model, it is possible to deduce some interesting consideration about the reduction of the Q-factor caused by metal and substrate losses. In effect the values of the series resistance (R_S) extracted with the described characterization procedure, are very close to the calculated ones on the basis of the technological data (sum of the metal squares multiplied for the square resistance R_{\square}). This allows to affirm that the dissipative effects due to the finite conductivity of the metals are totally described by this element; therefore the other resistors R_{SUB} , R_{SUBin} and R_{SUBout} model the substrate losses only. Then, if we define Q_{MET} as the quality factor of an inductor realized on an ideal substrate ($\rho_{\text{SUB}} \rightarrow \infty$) and with a real metal ($\rho_{\text{MET}} \neq 0$) and Q_{SUB} as the quality factor of an inductor realized on a real substrate ($\rho_{\text{SUB}} < \infty$) and with an ideal metal ($\rho_{\text{MET}} \rightarrow 0$), we obtain two useful parameters to analyze separately the inductor's losses. The Q-value of the actual

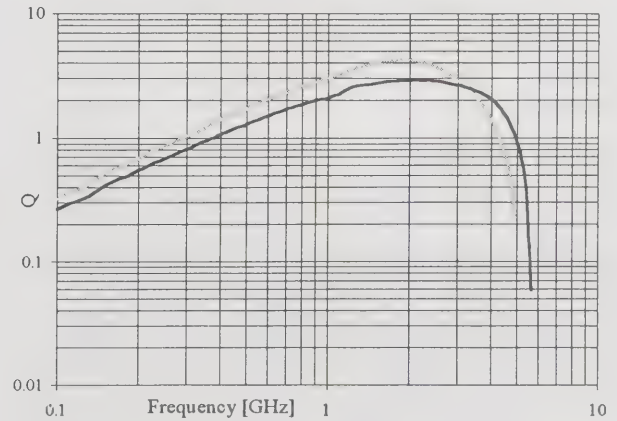


Fig. 6: Q versus frequency deduced from measurements of the integrated inductors; black line refers to CMOS inductor, gray line to BiCMOS inductor.

	L (nH)	R _S (Ω)	C _F (fF)	N _C	R _{SUB} (Ω)
Ind1 b	2	4.5	36	2.5	200
Ind2 b	2.4	6.9	39	2.5	180
Ind3 b	6.8	10.5	43	2	180
Ind4 b	13	20.5	58	2.3	200
Ind1 c	3.7	10	18	2	200
Ind2 c	7.3	15.5	22	2	180
Ind3 c	11.4	21	42	2	160

Table 2: extracted lumped parameters of the new model relative to the Y_S admittance.

inductor is given by the usual Q-factor composition rule:

$$\frac{1}{Q} = \frac{1}{Q_{MET}} + \frac{1}{Q_{SUB}} \quad (3)$$

We then apply this approach to the designed and characterized inductors (Tab. 1) focusing our attention in the two frequency typically used in commercial wireless applications; we obtain, for ind2_c, these values:

$$Q_{MET} = 3.16, Q_{SUB} = 13.83 \quad \text{at } f = 900 \text{ MHz}$$

and

$$Q_{MET} = 5.47, Q_{SUB} = 5.24 \quad \text{at } f = 1.8 \text{ GHz}$$

These results indicate that metal losses are very important in the low spectrum of wireless communications, but at 1.8 GHz their contribution in the Q degradation is comparable to substrate losses.

Other considerations about the Q of the inductors can be deduced from the model. Fig. 7 reports a three-dimensional plot of the Q (at 1.8 GHz) as a function of the series resistance and of the substrate resistance. This allows to estimate how much is possible to increase the quality factor increasing the substrate resistance or reducing the metal series resistance.

Conclusions

An improved equivalent circuit for Si integrated inductors is presented. This model, characterized by the introduction of a substrate resistance magnetically coupled with the inductor, accurately describes the phenomenon of the inductive coupling of dissipative currents in the low-resistive CMOS substrate. This allows more accurate simulation about RF ICs, accurately describing the parasitic effects introduced by integrated inductors.

The proposed equivalent circuit allows to understand which are the most convenient ways to improve the value of the Q-factor.

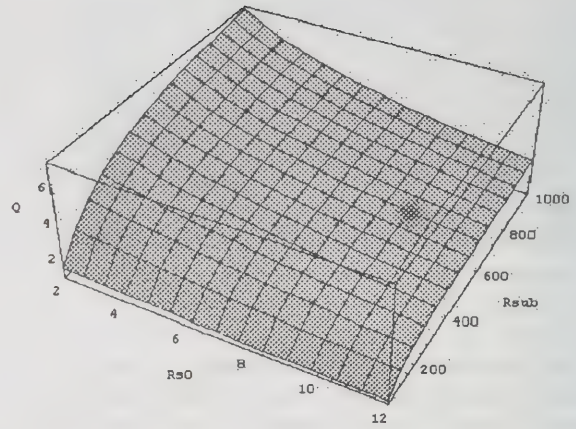


Fig. 7: quality factor at 1.8 GHz vs R_{S0} and R_{SUB} , based on the new model (Ind2_c inductor).

Acknowledgement

The authors wish to thank Dr. Ernesto Perea, of SGS-Thomson microelectronics, Grenoble (France), for the fruitful discussion about the development of this new equivalent model and the RF group of SGS-Thomson Microelectronics, Catania (Italy), in particular Dr. Privitera, for his help in the measurement of the integrated inductors.

References

- [1] C.Patrick Yue, S.Simon Wong, "On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC's", *IEEE Journal of Solid State Circuits*, vol. 33, N° 5, May 1998.
- [2] K. B. Ashby, I. A. Koulias, W. C. Finley, J. J. Bastek, and S. Moinian, "High Q Inductors for Wireless Applications in a Complementary Silicon Bipolar Process", *IEEE Journal of Solid-State Circuits*, vol. 31, N° 1, January 1996.
- [3] M.Soyuer, J.N.Burghartz, K.A.Jenkins, S. Ponnappalli, J.F.Ewen, and W.Pence, "Multilevel Monolithic Inductors in Silicon Technology", *Electronics Letters*, vol. 31, N° 5, March 1995.
- [4] N.M.Nguyen and R.G.Meyer, "Si IC-Compatible Inductors and LC Passive Filters", *IEEE Journal of Solid-State Circuits*, vol. 25, N° 4, August 1990.
- [5] H.M.Greenhouse, "Design of Planar Rectangular Microelectronic Inductors", *IEEE Trans. on Parts, Hybrids and Packaging*, vol. PHP-10, N° 2, June 1974.
- [6] P.Arcioni, R.Castello, G.De Astis, E.Sacchi and F.Svelto, "Design and Characterization of Silicon Integrated Inductors", *Proc. of IMTC 98*, S.Paul, MN, May 18-21 1998, vol. 2, pagg. 1395-1401.

Spiral Inductor Substrate Loss Modeling in Silicon RFICs

William B. Kuhn and Naveen K. Yanduru

Department of Electrical and Computer Engineering
261 Rathbone Hall, Kansas State University, Manhattan, KS, 66506
785-532-4649, 785-532-1188(fax), wkuhn@ksu.edu

Abstract

Spiral inductors constructed in Silicon IC technologies possess limited quality factors due to series resistive losses, and losses within the semiconducting substrate. A new model for the less understood substrate losses illustrates how these losses can be minimized, providing quality factor increases of up to 230 percent over un-optimized designs.

Key words: Spiral inductor, Optimization, Silicon, RFIC

1. BACKGROUND AND INTRODUCTION

Spiral inductors are finding increasing application in RF integrated circuits operating at cellular, PCS, and ISM frequencies from 0.8 to 2.4 GHz. Unfortunately, when constructed in standard, low-cost, silicon IC technologies, these inductors suffer from poor quality factors (Q), limiting the circuit performance obtained. The Q of on-chip inductors is ultimately limited by the series trace resistance of the thin metal layer in which the spiral pattern is formed [1]. However, energy losses from currents induced in the underlying substrate have also been found to be significant [2] [3], often limiting Q to 50 percent or less of the series- R limited value.

This paper presents a new lumped-element circuit model for analyzing substrate losses. Elements within the model are approximately derivable from basic physics, allowing the mechanisms behind the substrate losses to be better understood. A technique recently reported in [4] and [5] for minimizing substrate losses in standard process technologies is then studied in detail and results from the model are validated through electromagnetic simulations. The importance of considering the low bulk resistivity of the epi-wafers in CMOS/BiCMOS processes is also shown, and a new approach to producing high performance inductors in low-cost CMOS processes is proposed.

2. SUBSTRATE LOSS MODELING

Substrate losses result from I^2R losses associated with currents flowing within the semiconducting

silicon material below the inductor turns. As illustrated in Figure 1 these currents can be divided into two classes - those conducted through metal-to-substrate capacitances associated with individual turns of the spiral [1], and circulating eddy currents due to magnetic fields within the substrate [6].

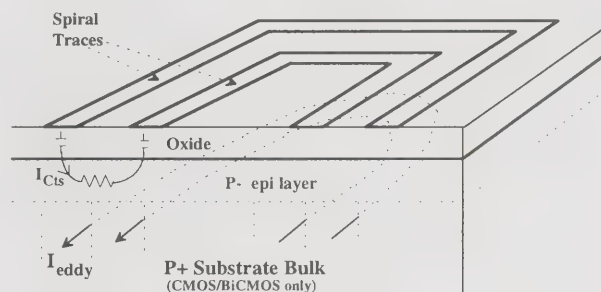


Figure 1. Spiral inductor substrate current flows.

An approximate, lumped-element circuit which models both effects can be created by modifying and extending those previously published in [7] and [8]. In Figure 2, L_i and R_i represent the inductance and series resistance of the i 'th spiral turn while C_{tt} represents interturn sidewall and fringing capacitance. The remaining elements model the substrate loss mechanisms described above. Capacitors C_{ts} represent turn-to-substrate capacitance while resistors R_{ts} represent connection of the bottom plate of these capacitances through the resistive substrate. (Precise modeling of these connecting resistances is complicated due to the 3-D bulk resistivity and current flow, but the circuit shown provides a reasonable approximation for many applications).

Finally, inductors L_{eddy} and associated resistors R_{eddy} model production of eddy current loops below each turn of the spiral. Here, the substrate is approximated as a 2-D resistive sheet, with the sheet broken into concentric, touching rings, one below each coil turn.

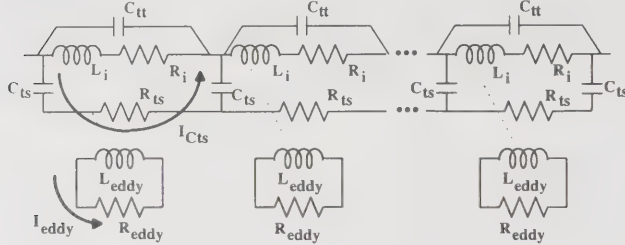


Figure 2. Simplified lumped element model of spiral and substrate losses. (inter-turn magnetic coupling not shown.)

All inductances within the model are interconnected through appropriate coupling coefficients when simulated, although only the coupling of the traces to the eddy current channels is shown. Capacitor values in the model are determined from standard area and fringe capacitance calculations, while resistor values are found from sheet resistance formulas after converting the substrate volume resistivity ρ into an approximate sheet resistivity. For R_{ts} , a rough approximation for effective thickness in which the currents are significant can be taken as equal to the inductor's turn-to-turn pitch p , yielding a sheet resistance of ρ/p , while a somewhat larger thickness on the order of one sixth to one third the inductor outer dimension is appropriate for use in approximating R_{eddy} .

3. ANALYSIS AND OPTIMIZATION

Assuming the substrate bulk resistivity is moderate (10 to 30 $\Omega\text{-cm}$), as in a typical bipolar fabrication process, the model of Figure 2 can be used to show that eddy currents (I_{eddy}) will be negligible and R_{eddy} and L_{eddy} can be removed from consideration. However, currents I_{Cts} flowing through R_{ts} will still contribute losses and are found to represent a major factor in lowering inductor Q .

Analysis of this loss mechanism suggests a technique for increasing Q by *reducing* R_{ts} , which can be accomplished in standard processes by placing a low resistance sheet below the coil turns. The use of an active N+ or P+ implant was suggested in [9], while the use of a poly, or patterned poly layer has been studied more recently in [4] and

[5]. In this section, the model of Figure 2 is used to explore the relationship between layer resistivity and inductor Q in more detail.

With sufficiently low resistivity, losses from currents conducted through R_{ts} can be reduced to negligible values, raising the overall inductor Q . However, if the resistivity of the new layer is too low, substantial eddy currents may begin to flow within it. Such currents will contribute new I^2R losses and begin lowering the inductor Q . These losses can be modeled with the L_{eddy} and R_{eddy} elements of Figure 2 in the same way that substrate eddy currents are modeled.

To illustrate the combined effects of decreasing losses in R_{ts} and increasing losses in R_{eddy} , Figure 3 shows simulated *parallel-resonant circuit* Q values obtained using the model of Figure 2. These results were obtained for a floating 8 turn, 16 nH inductor with 500 μm outer dimension, 0.03 Ω/\square metal, 2 μm oxide thickness, and varying values of sheet resistance below the coil turns. A significant local Q maximum can be seen between 20 and 100 Ohms/square where R_{ts} has fallen sufficiently low relative to X_{Cts} , but is still sufficiently high to prevent large eddy currents from forming.

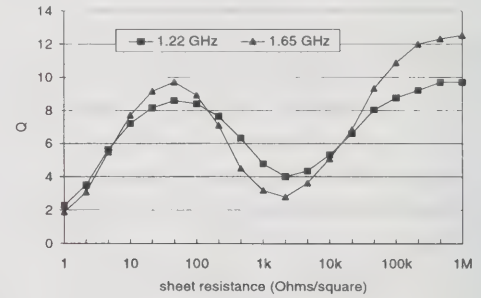


Figure 3. Q versus sheet resistance below coil simulated with model of Figure 2.

4. MODEL VALIDATION

To validate the results presented above, electromagnetic field-level analyses were performed with a commercial simulator. Results are shown in Figure 4 and Table 1 for various substrate constructions. For all simulations, a 7.5 turn floating inductor with a 500 μm outer dimension, 29 μm turn pitch, and 22 μm trace width was modeled and typical commercial process parameters were used: metal trace resistivity = 0.03 Ω/\square , trace-to-substrate oxide thickness = 2 μm , and substrate resistivity = 10 $\Omega\text{-cm}$ (except in the epi-wafer case). Inductor Q values listed were computed from the parallel equivalent resistance

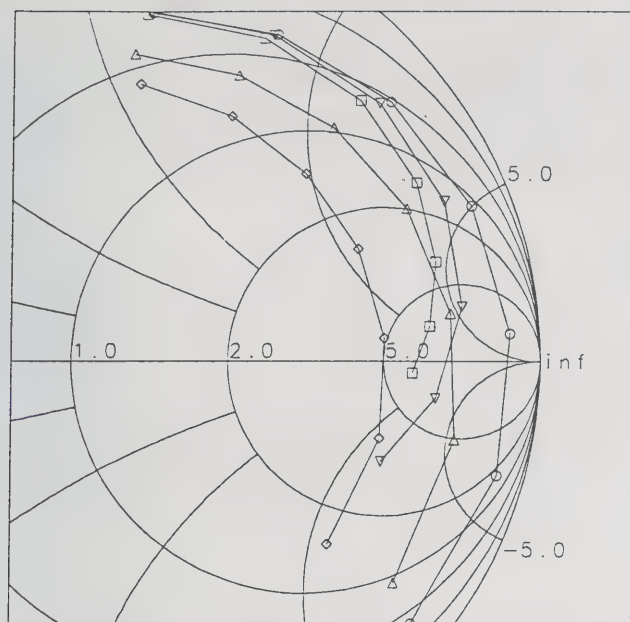


Figure 4. Inductor impedance at 0.5, 0.67, 0.91, 1.22, 1.65 2.23, and 3.0 GHz (clockwise from upper left) with 50Ω normalization

of the inductor's impedance divided by ωL at the operating frequency, equivalent to the definition of Q based on energy storage/dissipation in parallel resonant circuit applications. Similar behavior is obtained for series circuit Q values at lower frequencies, although the improvements and absolute Q values are reduced.

Table 1. Electromagnetic simulation results

Material below inductor	Q at	
	1.22 GHz	1.65 GHz
Substrate only	3.5	2.3
1 k Ω/\square well or channel-stop	4.7	3.0
25 Ω/\square sheet (N+, P+, or poly)	7.7	7.8
1 Ω/\square sheet (N+, P+, or poly)	1.6	1.3
10 Ω -cm epi over 0.01 Ω -cm bulk	2.9	2.7

Inductor Q is highest, as expected, for the case of 25 Ω/\square below the coil. Improvements over an undoped (10 Ω -cm) substrate and a substrate with a 1k Ω/\square well or channel-stop implant are 230% and 160% at 1.65 GHz (near self-resonance), and 120% and 65% at 1.2 GHz. Inductor Q is lowest when 1 Ω/\square is used indicating that this construction should be avoided. Q is also severely degraded by the low resistivity bulk of epi-wafers used in typical CMOS/BiCMOS processes to prevent latchup. This result has significant implications for designers considering the use of spirals in such fabrication technologies.

5. DECREASING SUBSTRATE LOSSES IN CMOS PROCESSES

The presence of a heavily doped (low resistance) substrate below the inductor turns in CMOS processes (and BiCMOS technologies derived from CMOS processes) represents a serious problem for inductor fabrication. One excellent solution pioneered by Chang, Abidi, and Gaitan involves selective removal of underlying substrate material through micromachining techniques [3]. Unfortunately, this technique adds cost to the overall chip, and may not be possible in some development projects due to a circuit designer's lack of influence over fabrication technology.

An alternative approach to solving this problem is suggested by the results presented in this paper. As illustrated in Section 3., it is acceptable to have a low resistivity layer below the inductor, provided the effective sheet resistivity of the new layer is greater than approximately 20 Ω/\square . Hence, a heavily doped $P+$ layer can be placed below CMOS circuits to prevent latchup and reduce crosstalk, provided the thickness of the layer is appropriately limited, or other measures are taken to limit the effective sheet resistance.

One possible construction is shown in Figure 5. In this case, a CMOS process wafer is constructed from a lightly doped bulk as in bipolar processes,

rather than the heavily doped bulk used in digital CMOS. A heavily doped, low resistance layer for latchup protection is then created on top of this bulk through epi-layer growth or other suitable fabrication techniques. Finally, the usual lightly doped epi-layer is formed, maintaining the same epi-layer thickness as in existing wafer constructions.

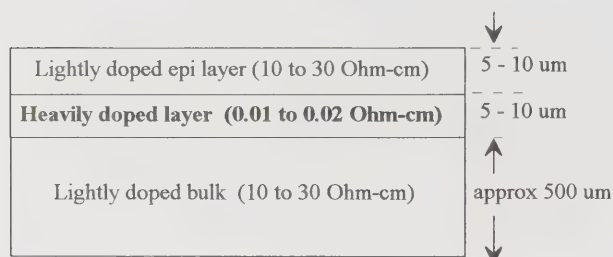


Figure 5. New epi-wafer structure which minimizes spiral inductor losses.

This new construction does have higher resistance to injected substrate currents from active devices than that of a traditional epi-wafer construction, since the doping is similar but the low resistivity region's thickness has been reduced. However, the difference is only significant for widely spaced devices on the die – an acceptable situation since latchup and crosstalk are most severe for devices with small separations (e.g. $\leq 10\mu\text{m}$). Over these distances, the new construction behaves substantially the same as in traditional wafers since the depth of current flow into the substrate is on the order of the separation distance.

6. CONCLUSIONS

A new model for approximating substrate losses in spiral inductors has been presented which illustrates the two independent loss mechanisms involved. Analysis of the model illustrates how inductor Q can be optimized in bipolar processes by placing a moderately low resistance layer beneath the inductor turns. The analysis also shows that it is possible to have a low resistance layer below CMOS circuits for latchup and crosstalk mitigation, without seriously degrading inductor quality factor. Application of these technologies can yield substantial improvements in overall inductor quality factor without resorting to modifications to the fabrication process.

7. ACKNOWLEDGMENTS

This material is based in part on work supported by the National Science Foundation under Grant No. EPS-9550487 and matching support from the state of Kansas. EM simulation software was provided by Hewlett Packard EEsof.

REFERENCES

- [1] Nguyen, N.M., and Meyer, R.G. "Si IC-compatible inductors and LC passive filters", IEEE J. Solid-State Circuits, 1990, SC-25, pp. 1028-1031
- [2] Ashby, K.B., Koullias, I.A., Finley, W.C., Bastek, J.J., and Moinian, S. "High Q inductors for wireless applications in a complementary silicon bipolar process", IEEE J. Solid-State Circuits, 1996, SC-31, pp. 4-9
- [3] Chang, J.Y.-C., Abidi, A.A., and Gaitan, M. "Large suspended inductors on silicon and their use in a 2- μm CMOS RF amplifier", IEEE Electron Device Lett., 1993, 14, (5), pp. 246-248
- [4] Yue, P. C., and Wong, S. S., "On-chip spiral inductors with patterned ground shields for Si-Based RF IC's", IEEE Symposium on VLSI Circuits, pp. 85-86, 1997.
- [5] Yue, P. C., and Wong, S. S., "On-Chip Spiral Inductors with Patterned Ground Shields for Si-Based RF IC's," IEEE J. Solid-State Circuits, Vol. 33, No. 5, pp. 743-752, May, 1998.
- [6] Craninckx, J., and Steyaert, M. S. J. "A 1.8-GHz low-phase-noise CMOS VCO using optimized hollow spiral inductors", IEEE J. Solid-State Circuits, 1997, SC-32, pp. 736-744
- [7] Long, J.R., and Copeland, M.A. "The modeling, characterization, and design of monolithic inductors for silicon RF IC's", IEEE J. Solid-State Circuits, 1997, SC-32, pp. 357-369
- [8] Kuhn, W. B., Elshabini-Riad, A., and Stephenson, F.W. "Centre-tapped spiral inductors for monolithic bandpass filters", Electron. Lett., 1995, 31, (8), pp. 625-626
- [9] Kuhn, W. B., Stephenson, F.W., and Elshabini-Riad, A., "A 200 MHz CMOS Q-Enhanced LC Bandpass Filter," IEEE J. Solid-State Circuits, pp. 1112 - 1122, August 1996.

Application of RF LDMOS Power Transistors for 2.2GHz Wideband-CDMA

Alan Wood and Warren Brakensiek
Wireless Infrastructure Systems Division,
Semiconductor Product Sector, Motorola, Inc.,
5005 East McDowell Road, Phoenix, AZ 85008
Tel. (602)-244-4740
Fax (602) 244-4471
Email: RJRT20@email.sps.mot.com

Abstract

Wideband CDMA was conceived to carry high bit rate data and video information in addition to voice traffic. The IMT2000 standard has adopted the 2.11 to 2.17 GHz band for this service, posing the design challenges of higher operating frequency band than the PCS 1.8-1.99GHz and wider modulation bandwidth for the signals. This paper describes the results for a 60 Watt single-ended and a 120 Watt push-pull RF LDMOS device characterized with W-CDMA signals in the IMT2000 frequency band.

Introduction

Third generation cellular systems are currently undergoing trials in Japan. Wide-band Code Division Multiple Access (W-CDMA) has been chosen to handle the wider bandwidth signals necessary to provide a variety of services including voice, data (384kB/sec) and video (2MB/s) signal transmission [1]. The frequency band 2.11 through 2.17 GHz has been allocated for this service and, depending on information bandwidth requirements, the spreading code is 4.096, 8.182, or 16.384Mbits/sec. Occupied channel bandwidth is 5, 10 and 20MHz respectively. System requirements demand high linearity because of the adjacent channel power ratio (ACPR) interim specification (-55dBc) and the need for up to 4 carriers to be amplified simultaneously by the base station RF power amplifier. Meeting these exacting specifications requires either a highly linear low efficiency class-A or a feed-forward error corrected class-AB amplifier. RF LDMOS can fulfil both these roles although the class-AB solution is preferred for its higher efficiency, despite the added system complexity.

This paper characterizes the performance of RF LDMOS technology for W-CDMA applications, demonstrating that excellent power gain and linearity can be obtained. Because of the low-Q internal matching networks and high performance of RF LDMOS operation in the 2.1-2.2GHz frequency band is readily achievable for RF transistors nominally designed for PCS applications. Special circuit techniques are necessary to handle both multi-carrier and wideband modulated signals. For instance, gain flatness and phase flatness across the frequency band is extremely important so that these circuits are usable in

a feed-forward error corrected power amplifier. Saturated power and gain has been sacrificed to achieve these desirable attributes.

Device description

RF LDMOS is a modified n-channel MOSFET specifically designed for both high frequency operation and power amplifier applications[2]. A sinker connects the source to the backside of the die. Gate and drain connections are made to the topside with wirebonds. Fig. 1 is a photograph of the 2GHz, 60 Watt rated RF power transistor. This transistor has a single section of input matching consisting of a shunt MOS-capacitor and series wirebonds that constitute the inductance of the low-pass matching and impedance transforming network. The matching network has relatively low-Q which enables the transistor to be operated over the frequency range 1.8 - 2.4GHz. Both the MOS-capacitor and the two transistor die are eutectically bonded to the package flange to provide a low inductance source ground path and low thermal resistance.

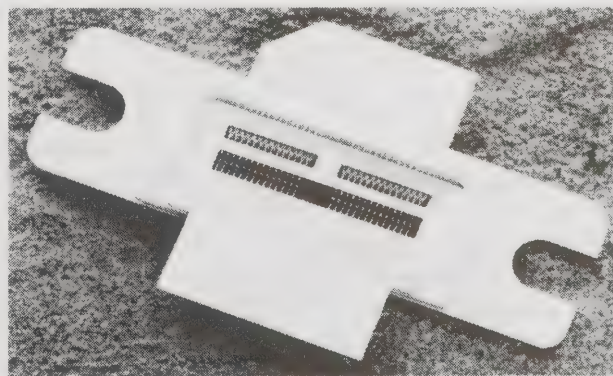


Figure 1. Photograph of 60W RF LDMOS device

The 120 Watt push-pull transistor is shown in Fig. 2. This transistor is similar in construction to the 60 Watt device, with each half of the push-pull consisting of two transistor die and a single MOS-capacitor. Both are mounted on a common flange.

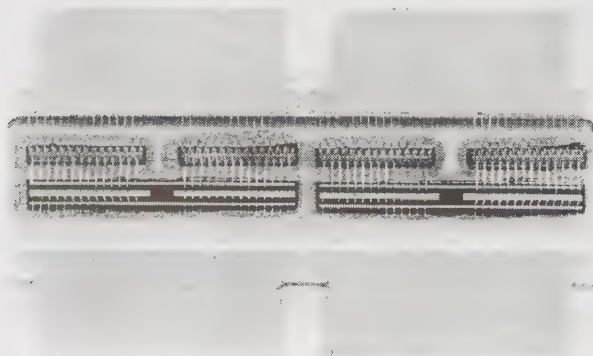


Figure 2, Photograph of 120W Push-Pull Transistor

Both transistors are ideally suited for the output stages of a RF base station power amplifier for W-CDMA applications.

Circuit Considerations

There are two major circuit design goals; a) matching the device across the frequency band 2.11 to 2.17 GHz while compensating for increased gain at the low end of the band by selectively mis-matching the output, b) designing the bias network such that it has relatively high impedance at the operating frequency but very low impedance at least up to the maximum modulation frequency or maximum carrier separation (60MHz). Low impedance also needs to be maintained up to several hundred MHz, for oscillation suppression over frequencies where the device has very high gain. Broadband performance for the nominally rated 60 Watt single-ended LDMOS transistor[3] is shown in Fig. 3

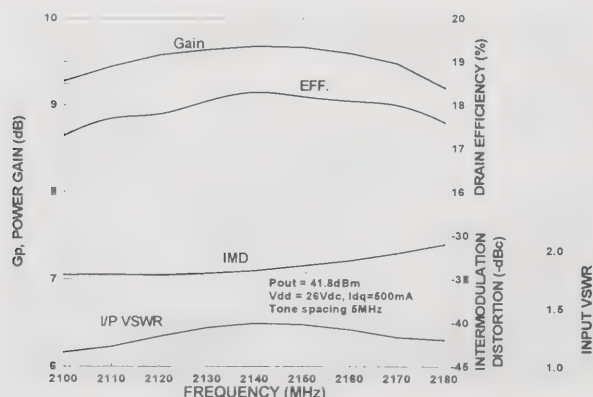


Figure 3, 2.1-2.2GHz Broadband Performance of 60 Watt RF LDMOS transistor.

Peak to peak gain flatness is 0.35 dB in the frequency range 2110 MHz to 2170 MHz. Minimum gain is 9.7dB. Input VSWR is lower than 1.6:1, and at the 41.8dBm power level intermodulation distortion is less than -30dBc. Electrical delay is 4.52nSec and deviation from linear phase is 1.25° peak to peak [Fig. 4].

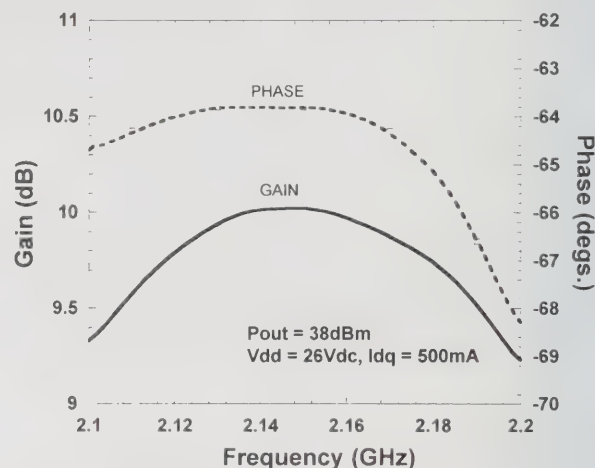


Figure 4. Gain and Phase versus Frequency for 60 Watt RF LDMOS

Power gain and insertion phase are shown in Fig. 5 as a function of output power. At 2.16GHz P1dB is 46dBm. Insertion phase can be seen to decrease as the gain compresses.

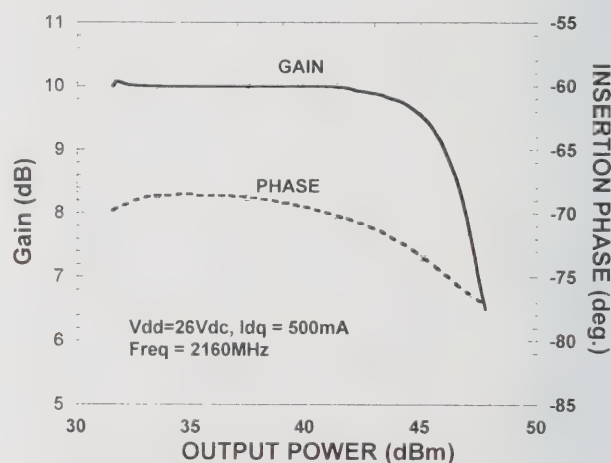


Figure 5, Gain & Phase versus Output Power

Push-pull transistors offer a 4x increase in device termination impedances for the same power level, which enables them to be matched to 50 Ohms over a broader bandwidth. This improvement in bandwidth is illustrated in Fig. 6 for a nominally rated 120 Watt push-pull transistor [4] consisting of two 60 Watt devices on a common flange. Gain flatness is +/- 0.1dB over the 2.11-2.17GHz band with a minimum gain of 11.1dB at 44.8dBm output power.

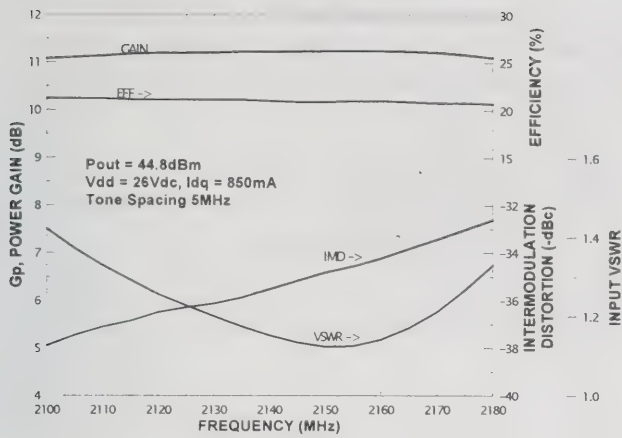


Figure 6, 2.1-2.2 GHz Broadband Performance for 120 Watt Push-pull Amplifier.

Two-tone intermodulation distortion is < -33 dBc at 44.8 dBm over this band and drain efficiency is $> 21\%$. Input VSWR is $< 1.4:1$ over the frequency range 2.1-2.2 GHz. Gain and phase as a function of frequency, with the push-pull transistor operated in the linear region (41 dBm), is shown in Fig. 7. Minimum gain between 2.11 and 2.17 GHz is 11.1 dB and peak to peak gain flatness is < 0.35 dB. Electrical delay for this circuit is 4.39 nsec with a deviation from linear phase of 920 psec peak to peak or 0.65° . For best error correction, both gain flatness and deviation from linear phase need to be minimized to maximize the amount of error correction that can be achieved in a feed-forward amplifier.

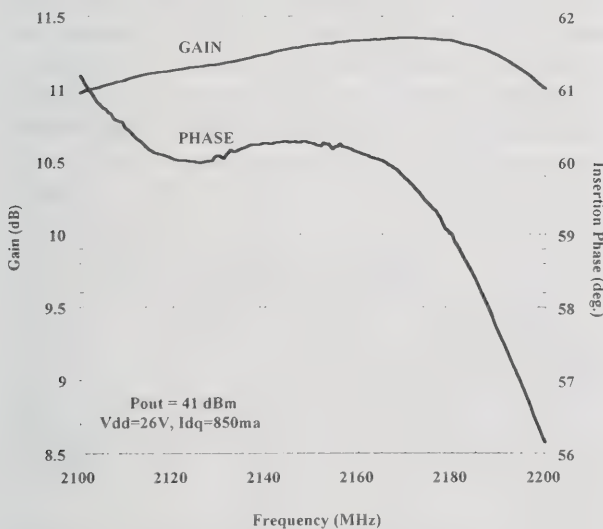


Figure 7, Gain & Phase versus Frequency for 120 Watt Transistor

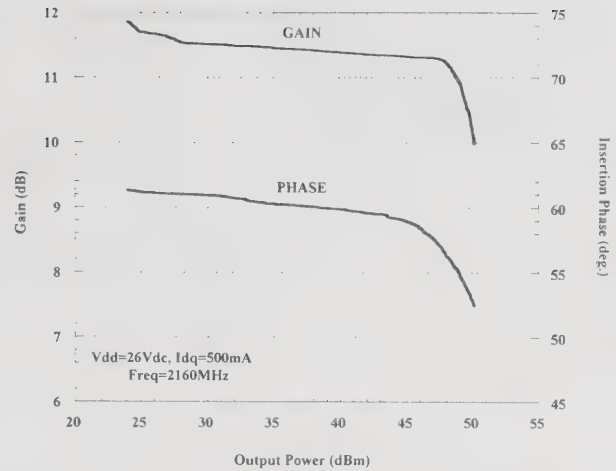


Figure 8, Gain & Phase versus Output Power

Insertion phase and gain have also been characterized as a function of output power at 2.16 GHz (Fig. 8). P1dB power is 49 dBm.

Linearity Performance

Two-tone intermodulation distortion, for the 60 Watt LDMOS circuit as a function of tone spacing is illustrated in Fig. 9.

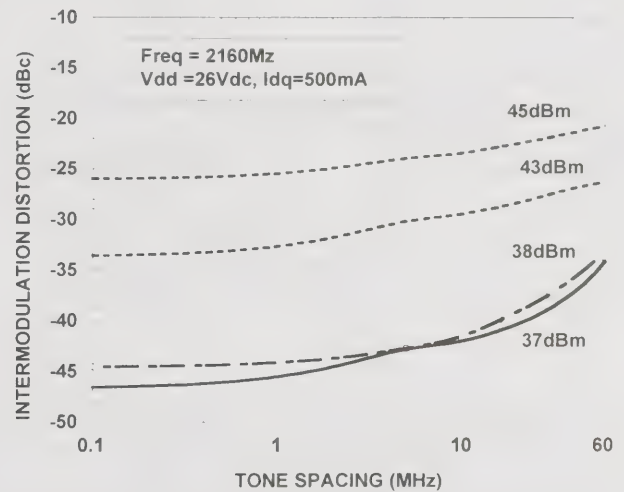


Figure 9, Intermodulation Distortion versus Tone Spacing

Bias networks are optimized to ensure minimum degradation at a 20 MHz tone spacing. Linearity is shown as a function of tone spacing in Fig. 9. Low IMD extends to 60 MHz tone spacing at output powers up to 43 dBm. ACPR performance with the 60 Watt device driven with a W-CDMA signal is shown in Fig. 10. This signal is spread with a 4.096 Mcps/sec spreading code. 50 channels are activated and the peak to average power or crest factor is 11 dB.

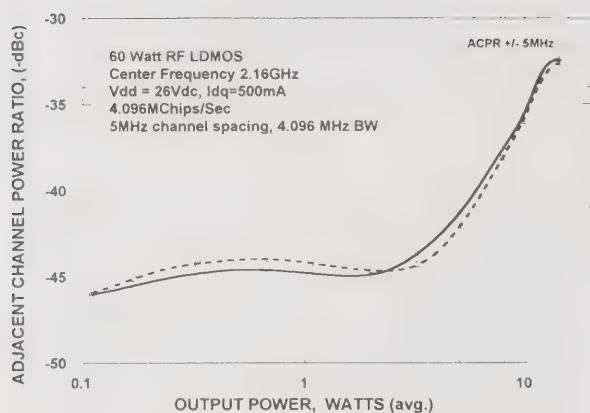


Figure 10, W-CDMA ACPR versus Output Power (60W RF LDMOS Transistor)

ACPR versus output power for a wideband noise signal that mimics a W-CDMA signal is illustrated in Fig. 11

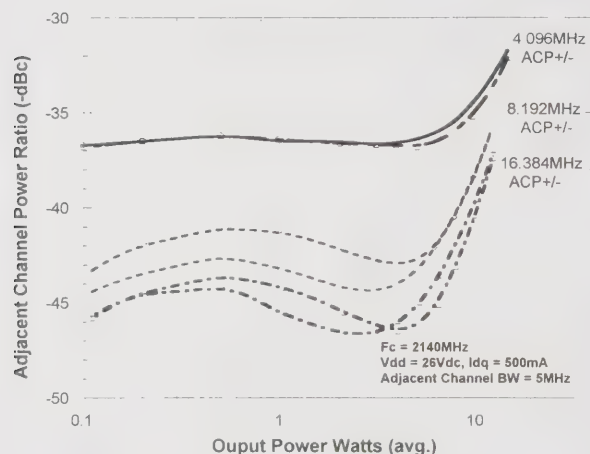


Figure 11, ACPR - Band-Limited Noise Signal

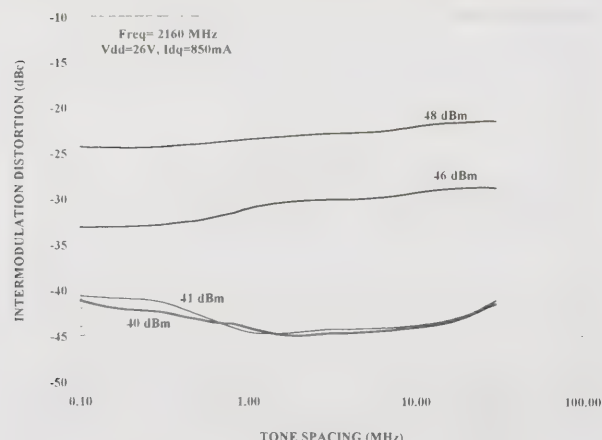


Figure 12, IMD versus Tone Spacing.

Intermodulation distortion versus tone spacing at different output powers, for the 120 Watt transistor, is illustrated in Fig. 12. At an output power of 41dBm

IMD is less than -40dBc. Fig. 12 shows the performance for a W-CDMA signal with a 4.096Mchips/sec spreading code. ACPR is > -40 dBc at 10 Watts average power.

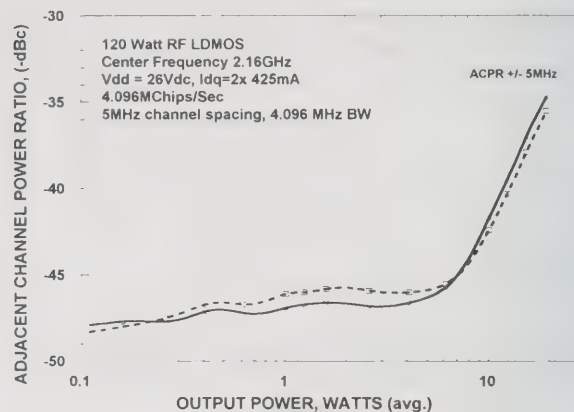


Figure 13, W-CDMA ACPR versus Output Power (120W Push-Pull Transistor)

Summary

With good circuit design and attention to bandwidth in the bias networks, RF LDMOS technology can provide excellent performance in the IMT2000 frequency band. Power gain in excess of 9dB can be achieved with good linearity at up to 60 MHz tone spacing. Gain flatness of +/- 0.2dB and phase flatness < 2° peak to peak make these circuits ideal for the output stages of feed-forward error corrected amplifiers. RF LDMOS, by virtue of the positive temperature co-efficient of resistance in the channel and high drain-source breakdown voltage, has excellent ruggedness characteristics into an output mismatch VSWR > 10:1, without any degradation in performance. Although, not characterized in this paper, RF LDMOS transistors are also ideally suited to class-A operation in the error amplifier of a feed-forward corrected base station transmitter power

References

- [1] Adachi, F. & Sawahashi, M, "Wideband Multi-rate DS-CDMA for Next Generation Mobile Communications Systems," IEEE Wireless Communication Conference, August 1997.
- [2] Camilleri, N et al; "Silicon MOSFETs, "The Microwave Device Technology for the 90's"; MTT-S International Microwave Symposium Digest 1993; pp. 545-548.
- [3] Wood A, Dragon A. & Burger, W., "High Performance Silicon LDMOS Technology for 2GHz RF Power Amplifier Applications," IEDM Tech. Digest 1996, pp. 87-90.
- [4] Wood, A et al, "120 Watt, 2GHz, Si LDMOS RF Power Transistor for PCS Base Station Applications", MTT-Symposium, June 1998.

Highly Efficient Linear Power Amplifier For 3.5 Volt NADC Cellular Applications

J. Staudinger, G. Norris, R. Sherman, G. Sadowiczak
Motorola, Semiconductor Products Sector
2100 E. Elliot Road, MS EL-712, Tempe, AZ 85284

Abstract

A highly efficient linear power amplifier has been developed for NADC mobile radio telephones operating with 3.5 volt batteries. The amplifier consists of two cascaded gain stages to achieve 25 dB gain at greater than 1 W output power with a corresponding power added efficiency of 50% at IS-136 specified ACPR linearity levels. The amplifier is realized using GaAs PHEMT devices to provide superior performance in optimizing the efficiency/linearity trade-off for this digital modulation format. Size of the amplifier is minimized by placing most of the circuitry on a single GaAs IC. An exception is the output matching circuitry where high Q lumped elements are selected to minimize loss. The circuit topology is chosen based on a design methodology where a large signal model of the PHEMT device is combined with behavior analysis techniques such that performance of the amplifier is simulated for a $\pi/4$ DQPSK modulated signal. This approach proved successful in sizing the FET devices as well as examining the performance improvements offered in using selected fundamental and harmonic terminations.

Key words: Power amplifiers, digital modulation, linearity

Introduction

Rapid growth of digital cellular and wireless systems continues throughout the world. These systems place demanding requirements on the power amplifier contained in the transmitter of the mobile unit, emphasizing cost, size, and performance at low voltage operation. Several of the cellular digital systems [1-3], including North American Digital Cellular (NADC), have adopted modulation formats which require a linear rather than saturated power amplifier. Since the efficiency of the amplifier has a direct impact on battery life and available talk time of the mobile radio, it becomes crucial to optimize the inherent linearity/efficiency trade-off in the circuit design process whereby the amplifier exceeds the specified linearity requirements while achieving the highest efficiency possible. In this paper, we present the results of a prototype linear power amplifier targeting NADC mobile phones operating with a nominal 3.5 volt supply. The circuit has been designed with the intent of maximizing efficiency while simultaneously exceeding the specified system linearity requirements.

Amplifier Performance Goals

Design goals for the amplifier are based on the NADC cellular standard (IS-136). The system uses a

$\pi/4$ DQPSK modulation format with a baseband root raised cosine filter with a bandwidth factor α of 0.35. Since the data rate is 48.6 Kb/s, the signal is band limited to 32.805 KHz. Linearity is specified in terms of both adjacent and alternate channel power ratio's (ACPR) at frequency offsets of ± 30 KHz and ± 60 KHz, respectively. Design goals for the amplifier are listed in Table 1.

Table 1
Amplifier Design Goals

Output Power	> 1 W
Gain	> 25 dB
Efficiency	> 45%
Supply Voltage	3.5 V
ACPR (adj)	> 26 dB
ACPR (alt)	> 45 dB
Input RL	> 10 dB
Frequency	824 - 849 MHz

PHEMT Device Technology

The amplifier is fabricated in Motorola's PHEMT-1 process, a 0.8 μ m PHEMT technology featuring an enhanced cross section refractory AlCu gate electrode defined by a passivation dielectric via in a single recess. The planar doped AlGaAs/InGaAs active

layer provides over 225mA/mm and 17V gate-drain breakdown for good efficiency and power density with 3.5V supplies. Pulsed I-V and on-wafer load pull data on these devices reveal minimal low frequency dispersion due to refinements in the recess and gate definition process. This allows for a high degree of accuracy in the large signal models used in the design process described later. The process is fully integrated and provides N+ resistors, MIM capacitors, and spiral inductors with 2 layers of metalization and 4 mil substrates. Although not used in this design, substrate thru vias are available to reduce source inductance in the large active devices required by the low operating voltage.

Circuit Design & Simulation

The amplifier is designed entirely using simulation tools. Particular emphasis was placed on examining the effects of device size, quiescent bias point, and harmonic source/load terminations to optimize the efficiency/linearity trade-off. In a multi-stage linear power amplifier, the design of each stage becomes intertwined with the others due to interaction of multiple active devices and interstage load lines. Therefore, the use of conventional load pull characterization methods to experimentally determine optimal conditions such as device size, load impedance terminations; bias ... was not pursued. Instead, a design methodology based on a large signal model of the PHEMT device coupled with harmonic balance and behavioral modeling methods was adopted.

A large signal model for a unit sized PHEMT device was developed. The device was first characterized by measuring S-parameters over a wide region of its I-V space using on-wafer probe methods. These measurements along with forward biased cold chip S-parameter measurements were used to determine extrinsic parasitic resistances, inductances, and intrinsic parameters for the large signal model.

The amplifier is designed by first considering the output gain stage which consists of the device and input/output matching circuitry (described by the fundamental, 2nd, and 3rd harmonic response) as illustrated in Fig. 1. A source inductance is included to model the use of bond wires connecting the source terminal to ground. The large output device (Q2) is represented by scaling parameters in the large signal model. Based on this circuit structure, load pull simulations are performed for a $\pi/4$ DQPSK stimulus with the goal of optimizing the linearity/efficiency trade-off. Harmonic balance simulations are first performed to determine the amplifiers am-am, am-

pm, and average supply current (I_{DD}) - input power (P_{in}) response when driven by a single tone CW signal at fundamental frequency f_0 . These simulations are repeated as a function of device sizing (gate width), quiescent point, and fundamental/harmonic source/load terminations. Using behavioral analysis techniques based on envelope transfer functions which are derived from the am-am, am-pm and I_{DD} - P_{in} data, the performance of the amplifier is simulated using a $\pi/4$ DQPSK stimulus compliant to the NADC standard [4-5]. This technique allows predicting output power, gain, linearity (adjacent, alternate, and 2nd alternate ACPR), and efficiency of the amplifier when driven with this digitally modulated stimulus.

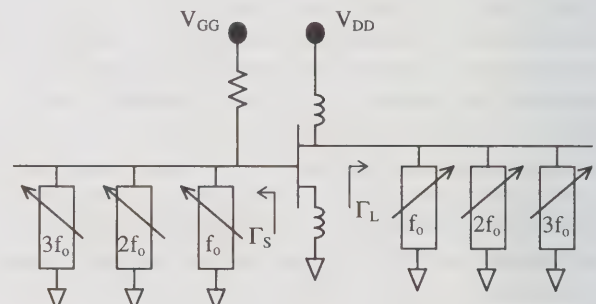


Figure 1 Load pull simulations were performed on the output gain stage as a function of bias and source/load terminations.

Using this simulation, the output load line was selected by varying the reflection coefficient (Γ_L) at the fundamental frequency (f_0). The results suggest an optimal Γ_L of 0.93 @ -176° and device gate width for Q2 of ~ 25 mm. This results in a power gain of greater than 12 dB (at an average output power of 31.5 dBm). In addition to the fundamental frequency termination, previous work [6] suggested amplifier linearity and efficiency can be improved with harmonic source & load terminations. The results of simulations in varying the angle of Γ_L at $2f_0$ are illustrated in Figure 2. In this case, the magnitude of the reflection coefficient is maintained at 0.98. These results suggest that for this circuit, where the output FET device utilizes several bond wires connecting the FET source terminal to ground, lower alternate ACPR is observed to occur an angle slightly greater than π . Additionally, the output power and gain are optimal near this angle as well.

The effect of higher order harmonic terminations as well as source terminations was investigated as well. The results suggest that while additional performance improvements are possible, the resonant circuitry needed to implement these conditions would

significantly add to the size and cost of the die and were thus not pursued.

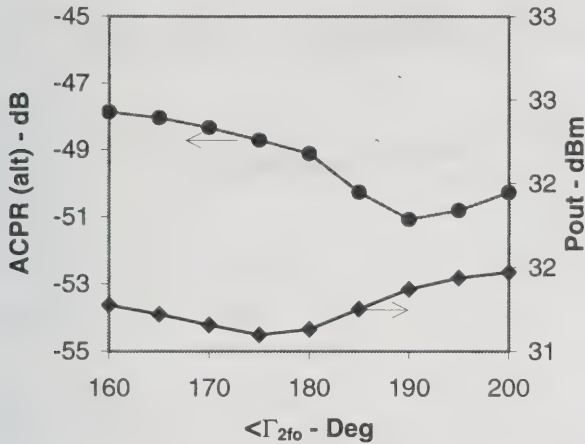


Figure 2. Effect of varying the angle of the load reflection coefficient at the second harmonic frequency. Results are obtained through simulation.

The design of the first gain stage was accomplished in much the same manner, although the design goals were somewhat different. Primary emphasis was placed on achieving reasonable efficiency and margin to linearity such that the linearity of the overall amplifier is set primarily by the second gain stage. The results of the design process are illustrated in Figure 3 which illustrates the linearity performance of the first gain stage as a function of drive level. Considering a second stage gain of 12 dB at an output power of 31.5 dBm, the first gain stage provides an output power of 20 dBm with more than 6 & 8 dB margin to the linearity specification, adjacent & alternate ACPR respectively.

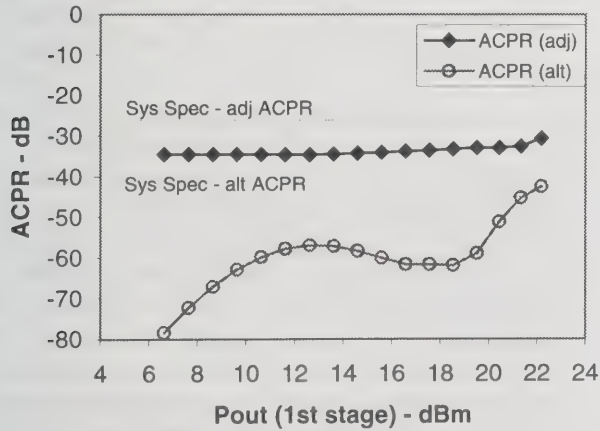


Figure 3. Linearity performance of the first gain stage obtained via simulation.

A schematic of the amplifier is depicted in Figure 4. As can be observed, most of the circuitry is contained on a single IC. An exception is the output matching structure and the Q1 drain bias decoupling element. On chip resonant circuits are utilized to set Γ_L at $2 f_0$ for both Q1 & Q2.

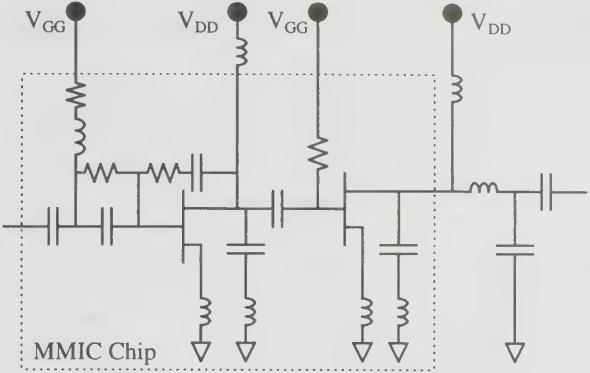
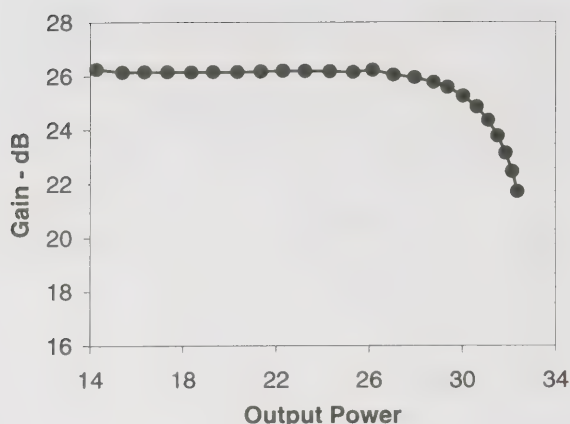


Figure 4. Schematic of amplifier.

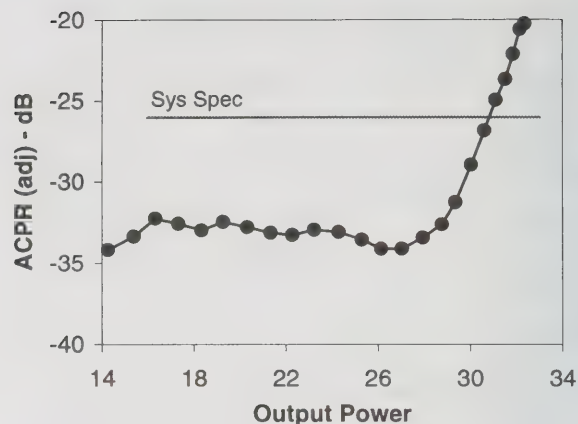
Fabricated Circuits & Measured Results

The circuit was on a 70 x 95 mil die. To test the RF performance of the NADC amplifier, the die was attached with conductive epoxy onto a hybrid printed circuit board. This test circuit board measures 1.5" x 1.5" and is fabricated from .015" FR4 material which is conductively laminated onto a piece of .0625" copper. The copper backing is used to provide a good thermal heat sink for the PA. All metalized traces are gold plated to facilitate device wire bonding. Multiple low inductance via holes are used to establish a top surface ground plane with each via hole measuring .015" in diameter. Surface mount 0812 capacitors are used for RF bypassing and smaller less parasitic 0603 capacitors are used in tandem with a series transmission line to set the required output load line.

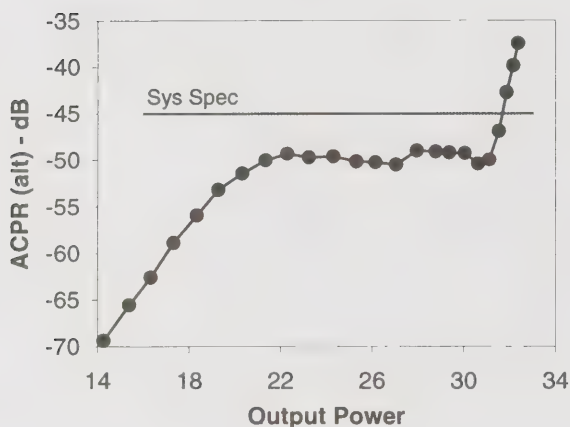
The performance of this amplifier when biased with a 3.5 volt supply and driven with a $\pi/4$ DQPSK NADC compliant signal is depicted in Figure 5. The amplifier achieves a small signal gain of greater than 26 dB. The amplifier linearity is illustrated by the adjacent and alternate channel power ratios (ACPR). Per the IS-136 system linearity specifications, the amplifier achieves 31.5 dBm output power with a corresponding power added efficiency of nearly 50%. Even at one dB output power back-off, giving additional margin to the linearity specification, the power added efficiency is about 45 %.



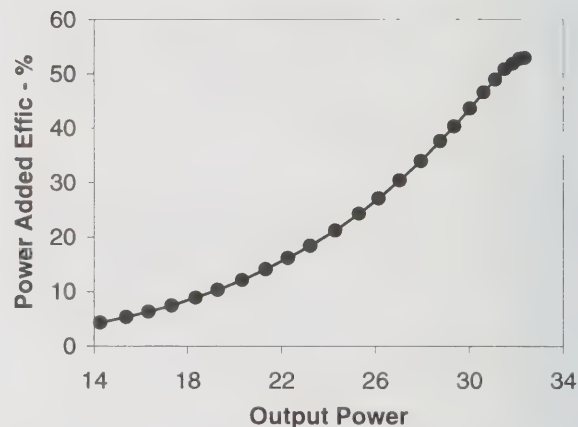
(a)



(b)



(c)



(d)

Figure 5. Measured amplifier performance when biased with a supply voltage of 3.5 volts and driven with a $\pi/4$ DQPSK NADC compliant signal. a) gain, b) Adjacent channel power ratio, c) Alternate channel power ratio, and d) Power added efficiency,

Summary

A high performance linear power amplifier based on GaAs PHEMT devices has been realized and demonstrated for application in NADC radio cellular handsets operating at 3.5 volt supplies. The amplifier consists of two cascaded gain stages to achieve 25 dB gain at greater than 1 W output power with a corresponding power added efficiency of 50% at IS-136 specified ACPR linearity levels.

References

- [1] TIA/EIA IS-54, "Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard," Telecommunications Industry Association, 1992.
- [2] TIA/EIA IS-95, "Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard," Telecommunications Industry Association, 1993.
- [3] Research and Development Center for Radio Systems (RCR), "Digital Cellular Communication Systems," RCR STD-27, April 1991.
- [4] A. R. Kaye, D. A. George, M. J. Eric, "Analysis and Compensation of Bandpass Nonlinearities," *IEEE Transactions on Communications*, vol. COM-19, October 1972, pp. 965-972.
- [5] J. Staudinger, "Applying The Quadrature Modeling Technique To Wireless Power Amplifiers", *Microwave Journal*, pp. 66-86, Nov., 1997.
- [6] J. Staudinger, G. Norris, "The Effect of Harmonic Load Terminations on RF Amplifier Linearity for Sinusoidal and $\pi/4$ DQPSK Stimuli", 1997 IEEE MTT-S International Topical Symposium on Technologies for Wireless Applications", pp 23-28, 2/23/97, Vancouver, BC, Canada.

A New Circuit Topology to Realize High Efficiency, High Linearity, and High Power Microwave Amplifiers

David M. Upton

Mayflower Communications Company, Inc., Billerica, MA
(508) 436-9600 x11, (508) 436-9988 (fax), dmu@mayflowercom.com

Peter R. Maloney

Raytheon Systems Company, Radar & Technology Programs, Sudbury, MA

Abstract

The unexpected and unprecedented growth in demand for increased communications capacity by telecommunications consumers has led to the implementation of a variety of multi-spectral digital modulation techniques. All of these have in common the need for improved amplifier linearity in order to be practical or to co-exist with other users in the same spectrum. Many systems also mandate that the amplifier efficiency be as high as possible to extend the operating time of portable or self-contained power sources. In addition to the above, many systems desire relatively high power outputs to be obtained for improving reliability or availability. A workable compromise is difficult to achieve for any given set of parameters. Much effort in the literature is devoted to designs of amplifiers that seek to meet the above requirements. In this paper, work that was reported previously is extended to a more general higher power case by introducing a previously unknown circuit topology variation of a Microwave Doherty Amplifier. The new topology shown easily extends the load line modulation principle to higher power designs while also providing a design technique that avoids many of the complications of two tightly coupled interactive stages. A momentary review of the previous work is included for completeness.

Key words: Power amplifier, linearity, efficiency, noise power ratio

Introduction

Doherty amplification was first developed during the vacuum tube era while high power and high efficiency transmitting techniques were being investigated.[1], [2], [3] Doherty amplification in vacuum tube and solid-state amplifiers use the technique of load-line modulation of one device by the other through a quarter-wave line. The circuit configuration of the classical Doherty requires quadrature drive to the inputs of a "carrier" amplifier and "peak" amplifier, respectively.

In operation, the control bias of these devices is offset so that the carrier amplifier operates near Class B while the peak amplifier operates near Class C. This results in the carrier amplifier providing the majority of the signal power while the peak amplifier cuts in when the carrier amplifier would otherwise be instantaneously saturated. The higher peak power achievable by this configuration can be explained by the fact that the load-line impedance of the carrier amplifier has been momentarily reduced by having the peak amplifier suddenly contribute current to the three-port junction at the other end of the quarter-wave line, while at the same time the peak amplifier also contributes output power to the load. The out-of-phase drive and the presence of the quarter-wave line are seen

as necessary conditions to enable the load-line modulation effect to occur. The offset biases simply maximize the efficiency obtained and reduce the imposed penalty of running the duplicated stage.

In the referenced papers and patent the authors explained the operating principles of the new Microwave Doherty Amplifier that had been realized using modern GaAs FET devices and transmission line techniques. [4], [5], [6] Operation at 1.3 GHz was shown with high efficiency and high linearity. In the second paper, more detail on evaluating the performance of the amplifiers under multi-tone conditions was included. The generation of amplifiers reported here has been evaluated using exactly the same techniques.

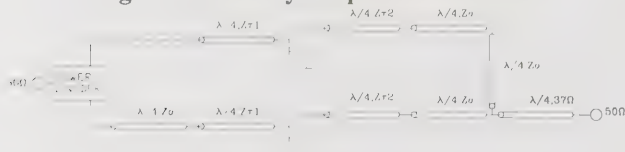
In the initial paper it was apparent that the reported technique was not directly applicable to considerably higher power designs since the realization of lower impedance lines on microwave substrate materials was not practical. The load-line impedance required for the active device scales inversely with power, all other conditions being equal. Thus, it is desired to find a method to employ load-line modulation techniques without working at the undesirably low impedances needed at the device level. It is also further desired to

remove the complications of designing such an amplifier with the tightly coupled interactions between the two stages and reduce the degree of difficulty in analysis to a more well understood circuit.

Description Of The New Topology And Practical Considerations

The basis of the new topology is to design two well-phase-matched class B amplifier stages that follow a well-known design procedure. Inputs to the two output stages are required to be 90 degrees out-of-phase so that the summation of power at the output will be performed correctly. In operation, the biases of the two stages will be offset slightly so that the “carrier” amplifier is the more active of the two. Then the outputs of these two stages, previously matched to provide 50 ohm output while presenting an optimum load-line to the device(s), are connected in a Doherty configuration, i.e., with the output of the so-called carrier amplifier through a quarter-wave line. The amplifiers are effectively connected in parallel through the output quarter-wave line so that these outputs combine at that point in a reactive power summation. Output may be taken from this point by matching the combined impedance (25 ohms) back to the desired 50 ohms. A conceptual schematic of the new configuration is shown in Fig. 1. A search of the literature shows that this technique is counter-intuitive as it was previously asserted that the “Doherty Effect” would not be practical once the tight coupling between the two stages was eliminated at the device level. Indeed, any additional matching between the quarter-wave line and the devices was seen to be disadvantageous as it is difficult to get the real part of the complex impedance to transform correctly with real lossy transmission lines or networks. [7]

Figure 1. Doherty Amplifier Schematic



Practical amplifiers may be designed and built using the technique outlined above using single-ended or push-pull Class B stages. There are no significant modifications required to a Class B stage for use in this amplifier configuration unlike the original tightly coupled circuit where each amplifier loaded the other one significantly as well as having challenging supply distribution and decoupling problems. These issues largely disappear with the new topology. The designer is free to implement the input and output matching to the devices in the conventional manner while the device load-lines are

similarly unconstrained. In operation, the apparent output load impedance will swing between 100 and 25 ohms at the carrier amplifier side of the quarter-wave line so designing for a 50 ohm load is reasonable. Provision for the input power split and 90 degree phase shift must be included in the design, of course, although these may be realized externally. The output quarter-wave line impedance is also not limited to 50 ohms. Since the individual stages are Class B amplifiers, techniques such as harmonic tuning of either the input or output matching networks or both are also permissible. Phase-matching requirements of the two amplifiers lend themselves readily to monolithic fabrication of the input circuitry and, in lower power cases where maximum efficiency is not a goal, the output circuitry.

Biasing of the two amplifiers is best accomplished by evaluating lots of devices on a normalized current density basis. That is, settling on an operating point for each device expressed by a current / emitter area or gate periphery. The drain/collector current of each device must be controlled to ensure that the device stays within its desired region of operation as the efficiency will suffer otherwise. It is unlikely that constant bias schemes will work well in this application, since the bias needs to adapt to increases in device temperature and signal characteristics. As in any Class B amplifier, the resistance to ground must be low to prevent self-biasing into the AB region. Broadband or multi-carrier modulation in the application of the amplifier will require particular attention to biasing and decoupling technique in order to extract the best performance.

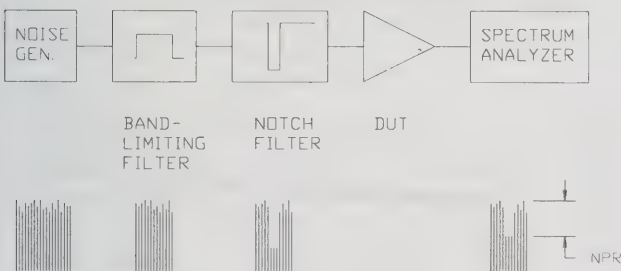
NPR Evaluation Method

As in the previous work, the amplifiers were measured using Noise Power Ratio techniques. This method is to be preferred over that of single tone CW or two-tone testing for several reasons. One of these is that such a driving signal provides a more accurate representation of amplifier characteristics when used with multiple modulated carriers such as multiple voice channels. Another is that the high peak-to-average ratio inherent in such excitation makes apparent subtle effects not easily observed with less complex signals. Finally, if the amplifier in question works well under these conditions, then a degree of confidence exists about its performance under less demanding circumstances.

The Noise Power Ratio method provides a practical means of observing low level non-linearities with multi-carrier systems with very low cost compared to alternative equipment. A broadband commercial noise source available from many vendors is amplified by low power small-signal amplifiers as shown in Fig. 2 and

bandlimited. This signal is then applied to a sharply tuned notch filter that places a notch in the applied spectrum corresponding to the loss of a single channel in a multiple channel system.

Figure 2. Noise Power Ratio Test Setup



Output from the amplifier under test is then observed on a spectrum analyzer paying particular attention to the definition of the notch as drive is increased. The NPR is then the depth of the notch at specified levels of output power. Low-level non-linearities in the amplifier being evaluated cause the notched region to fill up with broad-band noise due to high order intermodulation distortion. NPR is thus an excellent method to evaluate amplifiers that will see noise-like multi-carrier signals such as are common in modern digital communications.

Performance

Performance of the new amplifier topology can be seen in Figures 3 and 4, showing the DC Efficiency obtained versus the behavior of an ideal Class B, and the Noise Power Ratio (NPR) versus power output curve. The NPR signal is characteristic of a multi-tone (noise-like) signal of 10 MHz bandwidth. A Gain compression curve obtained during the NPR characterization is included in Figure 5. A modest amount of gain enhancement occurs that is indicative of Class B operation of the carrier amplifier. As an illustration of the low AM-PM of this amplifier configuration, the drain voltage was increased to 7.0 V with a CW drive applied. The change in insertion phase with drive level is shown in Figure 6. These conditions are representative of the output power obtainable with simple signals. A schematic of an amplifier designed with the concepts disclosed in this paper is in Figure 7.

Conclusion

A new variation to the microwave Doherty high efficiency amplifier technique allows scaling the amplifier to much higher output power. [8] In addition, the new technique illustrates that the load-line modulation principle may be extended beyond the reference plane of the active device through a matching network. Several

advantages result from being able to isolate the constituent amplifiers such as easier biasing, ability to employ harmonic tuning, and an almost unlimited increase in power handling capability. As the complexity of signals increases in communications applications, more emphasis will be needed on high efficiency highly linear amplifiers that allow active power output control for optimal power management. The new topology reported here appears to meet many of these requirements.

References

1. W.H. Doherty, "A new high efficiency power amplifier for modulated waves," Proc. IRE, Vol. 24, no. 9, pp. 1163-1182, September 1936.
2. F. E. Terman, Radio Engineer's Handbook, third edition, McGraw-Hill, 1943.
3. F. H. Raab, "High-Efficiency Amplification Techniques," IEEE Circuits and Systems, Vol. 7, no. 10, December 1975.
4. Upton, David M., and Maloney, Peter R. et al, "A Microwave Doherty Amplifier," IEEE MTT-S Symposium Digest, 1994.
5. Upton, David M., and Maloney, Peter R. et al, "A Microwave Doherty Amplifier," Applied Microwaves and RF, Winter 1994.
6. U.S. Patent # 5,420,541--A Microwave Doherty Amplifier.
7. F. H. Raab, "Efficiency of Doherty RF power amplifier systems," IEEE transactions on Broadcasting, Vol. BC-33, no. 3, September 1987.
8. U.S. Patent # 5, 568, 086--Pertaining to further improvements on the high-efficiency microwave RF power amplifier.

Figure 3. DC-RF Efficiency vs. Output Power

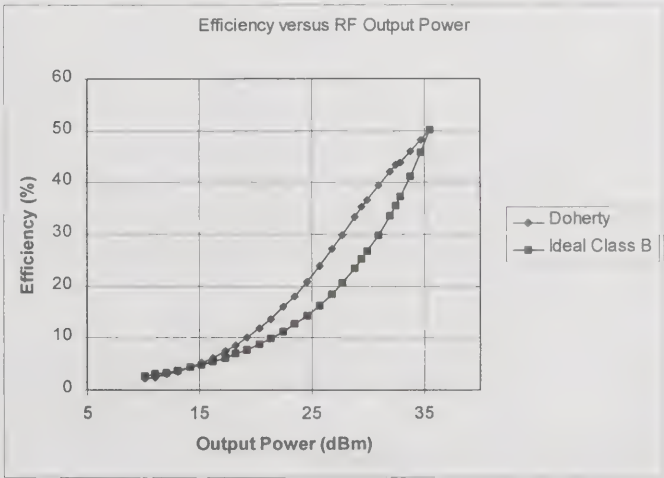


Figure 4. Noise Power Ratio vs. RF Power Out

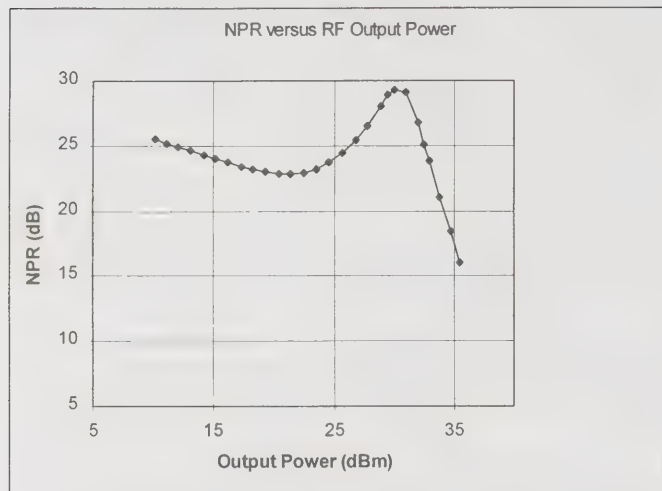


Figure 5. Gain vs. RF Output Power

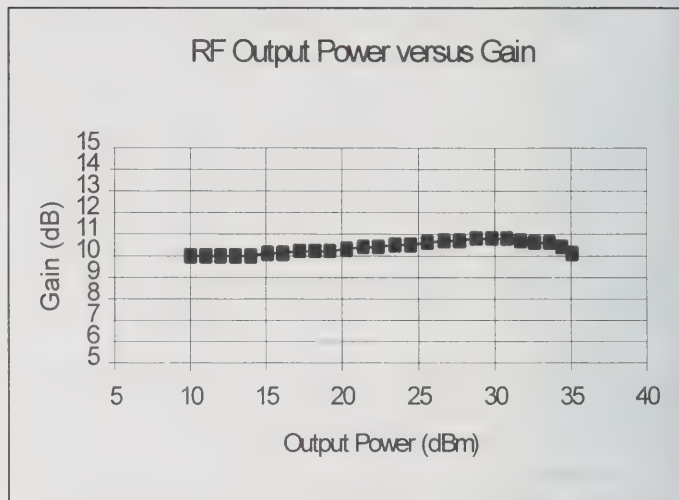


Figure 6. Insertion Phase vs. RF Output Power

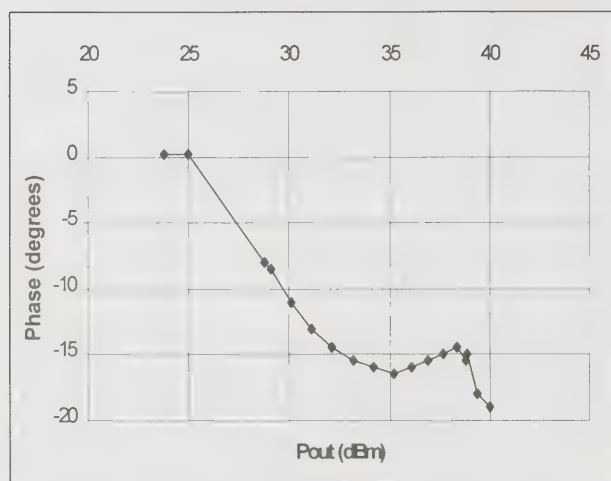
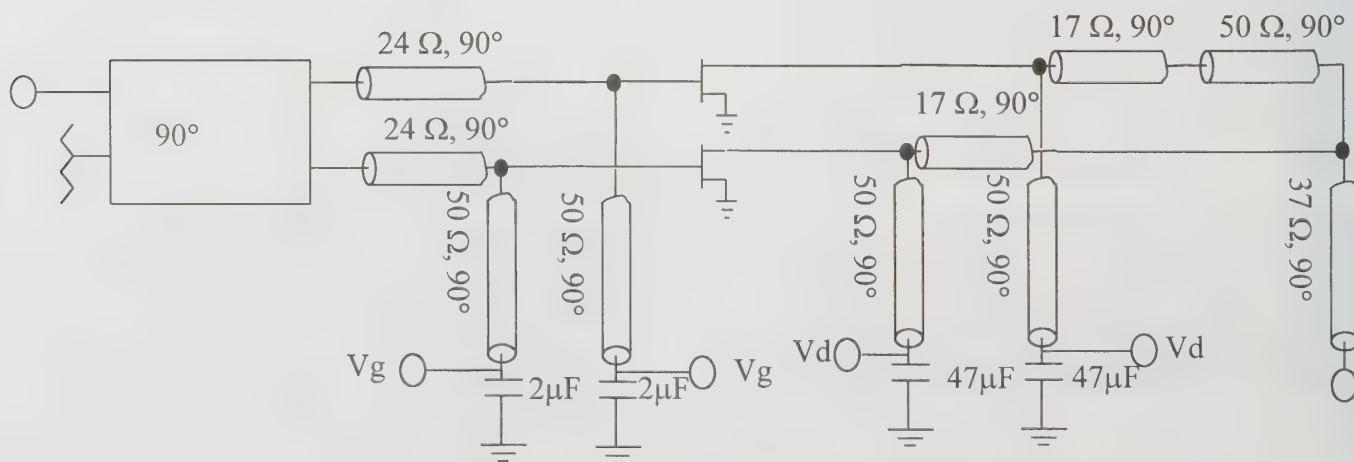


Figure 7. Detailed Schematic



GaAs HBT Power Amplifier with Smooth Gain Control Characteristics

Esko Järvinen

Nokia Research Center, P.O. Box 407, FIN-00045 Nokia Group, Finland

Email: esko.jarvinen@research.nokia.com

Abstract

A gain control scheme giving simultaneously an effective power switch-off operation and a smooth gain control characteristics is presented. The concept was applied to a 3-stage power amplifier, which was a part of a 900-MHz transmitter chip. The chip was realised by using GaAs MMIC technology and was mounted on a ceramic substrate. The power amplifier was tested separately by mounting it directly on a conventional FR4 board. In this lossy environment the power amplifier delivered an output power of 33.2 dBm with a supply voltage of 3.5 V. The maximum value of the gain control slope was 22 V/V showing a good agreement to the simulations.

Introduction

Wireless communications systems set a challenging task to the power amplifier designers when they try to fulfil all the specifications. In addition to the common and well understood requirements such as output power and efficiency there are also other properties, which are as important but easily underestimated. An important parameter, which has a broad effect on the system design, is the gain control slope of the power amplifier.

A common way to control the gain of the transistor stage is by adjusting its operating current [1, 2]. For the optimum performance the whole control voltage range (e.g. 0...2 V) should be utilised which is difficult, especially with an exponential-law component such as BJT or HBT.

A conventional resistive bias circuit is poor in this respect and the true control voltage range is only few hundred millivolts. A better bias structure, well suited for power amplifiers, is described in this presentation. The idea is based on the better utilisation of the voltages available in the mobile phone.

This concept is adopted in a 3-stage 900-MHz power amplifier realised by using GaAs HBT technology. The power amplifier is a part of the transmitter chip consisting of a modulator, a power amplifier, and an LO-generation circuit.

The chip will be mounted onto the ceramic substrate with integrated capacitors and resistors. This offers good thermal and electrical properties and reduces the total area of the power amplifier compared to a

conventional solution with a plastic package and numerous discrete components around it on a printed circuit board.

Bias Circuits

Fig. 1a shows a typical HBT stage with resistive biasing. This type of bias circuit reduces the control range, because the operation is controlled directly by the base voltage.

A better performance is achieved when using current mirror type of biasing (Fig. 1b), where the collector current I_{c1} is controlled by the control current I_{pc} . This circuit is also more stable against temperature variations than the previous one.

The performance can still be improved if we together with the control voltage utilise the regulated voltage V_{reg} available in the phone. Fig. 1c shows one possible solution, which is derived from the current mirror bias circuit.

Fig. 2. Shows the simulated collector current as a function of the control voltage V_{pc} for the circuits 1a - c. The circuits have been designed so that the maximum current, achieved at $V_{pc} = 2.2$ V, is the same for all the circuits. With the resistive bias circuit the whole operation from zero to maximum current is achieved within a V_{pc} -range of 0.6 V. The circuit in Fig. 1b shows a range of 1.0 V, and the widest range of 1.4 V, is achieved with the circuit of Fig. 1c.

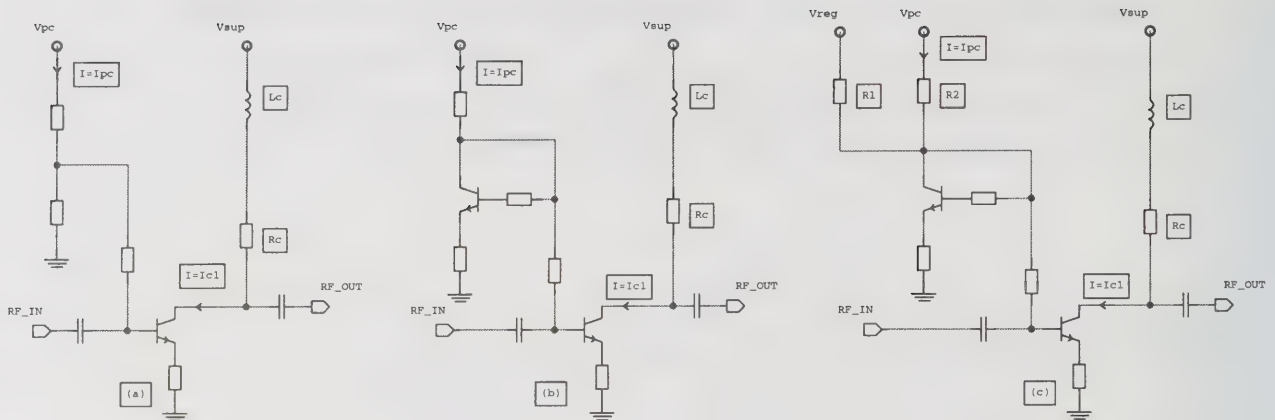


Fig. 1. Bias circuits. (a) Resistive (b) current mirror (c) current mirror with enhanced control range.

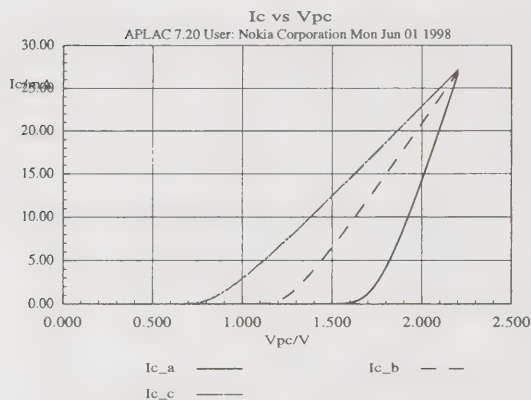


Fig. 2. Simulated I_c versus V_{pc} for the circuits in Fig. 1.

Power Amplifier Design

The schematic diagram of the power amplifier part of the chip is shown in Fig. 3. It is a three-stage HBT MMIC amplifier designed to fulfil GSM mobile requirements. The output stage consists of 48 custom designed unit cells. Simulations showed an output power capability of over 3 W with a power supply voltage of 3 V, in which case the optimum load resistance is around 1Ω . This makes the realisation of a low-loss output-matching circuit a challenging task. Therefore we designed the chip to be mounted on a ceramic substrate as a carrier providing good electrical and thermal properties. A low-pass topology was employed to the matching circuit and this reduces effectively the harmonic content at the output. For

improving the efficiency a second harmonic trap was used by utilising two bond wires from the total of ten at the output to form a series resonant circuit together with an off-chip capacitor.

The interstage matching between the last stage Q3 and the driver stage Q2 is even more critical than the output matching. The input impedance of the last stage is low, below 1Ω , and there is a high uncertainty in its value. The optimum load for the driver stage is around 10Ω depending on the gain of the last stage. It is possible to use a simple series-C, shunt-L topology for matching, but it leads to inconvenient parameter values (high C, low L). Furthermore, it is not possible to tune the impedance level. To get some freedom in design, the matching is therefore realised mostly off chip. Double pads at the base of the last stage and at the collector of the second stage give flexibility in choosing the topology of the interstage matching circuit.

The first stage employs a negative feedback to improve the stability and input matching. In this case a series-C, shunt-L interstage matching is a good choice. The inductance is formed by the bonding wire and the microstripline on the ceramic substrate. The first stage has a separate ground to avoid stability problems often caused by the common ground inductance.

The last stage transistor Q3 is biased to Class-AB by using a base current driving transistor Q7 to reduce the control current. From many possible temperature compensation circuits, the one realised here with Q6, provided the required behaviour. The gain of the amplifier is controlled by the bias of the first and second stages.

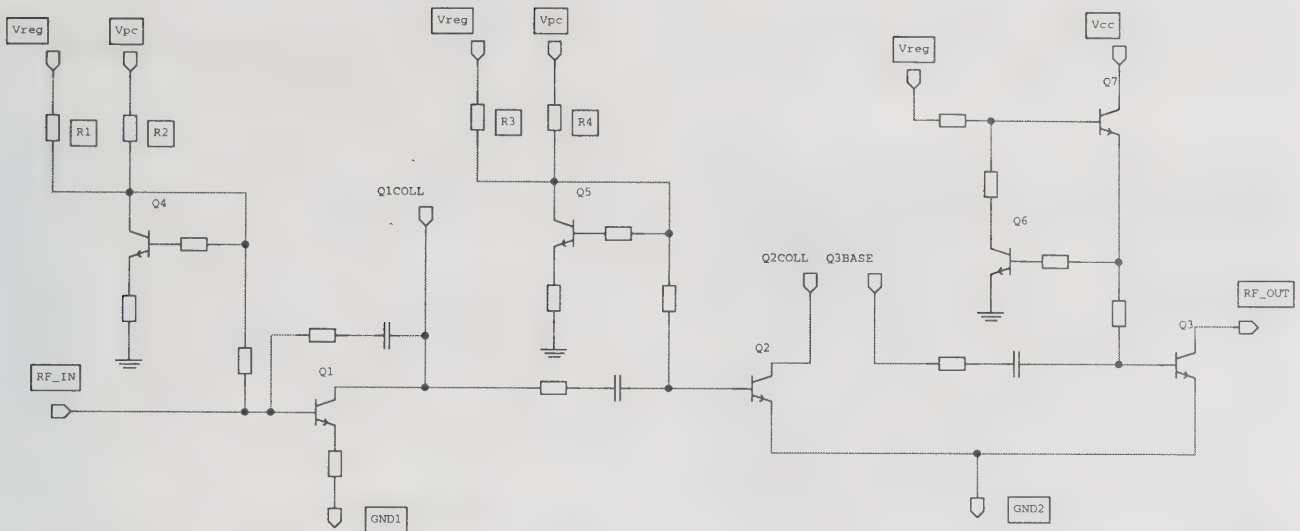


Fig. 3. Schematic diagram of the power amplifier chip.

There are two important aspects concerning the power control of the power amplifier. Firstly, the power should be easily switched off to a high isolation state. For a GSM power amplifier the recommended output power between the bursts is below -40 dBm which corresponds to a forward isolation of 40 dB. Secondly, the output power should be controllable below the actual control range. The nominal output power of a GSM transmitter is 0...35 dBm, but for the proper operation of the control loop the starting level at the beginning of the burst is in the range of -20 to -10 dBm.

For a typical circuit (Fig. 1a), where the switch-off and power-level-control functions are realised with a single control voltage V_{pc} , these requirements are contradictory and reduce the control voltage range available to the actual power control. In the bias scheme described in Fig. 1c we utilise the fact that it is possible to switch off also the regulated voltage V_{reg} between the bursts. When V_{pc} and V_{reg} are both at a low state, the amplifier is very effectively switched off with a high forward isolation. On the other hand, by a proper selection of resistors $R1$ and $R2$, it is possible to vary the threshold of the tuning curve and define the gain control range.

There is, however, one drawback concerning this circuit. V_{pc} has now to be able to both source and sink the control current. However, this is usually not a problem. On the other hand, the maximum I_{pc} (typically few milliamps) is now reduced from the value of a typical circuit.

An important parameter affecting the design of the control loop is the gain control slope

$$K_v = \Delta V_{out} / \Delta V_{pc},$$

where V_{out} is the peak RF voltage at a 50-Ω load and V_{pc} the control voltage. It is not unusual that a value higher than 50 V/V can occur when the bias control circuit of Fig. 1a is used. According to the simulations the maximum value of K_v in our case is below 21 V/V (Fig. 4), which makes the design of the control loop easier.

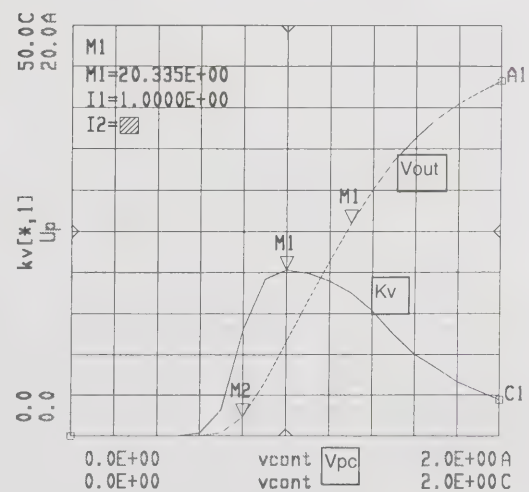


Fig. 4. Simulated RF peak voltage to a 50-Ω load and a gain control slope K_v of a three-stage amplifier as a function of the control voltage V_{pc} .

The chip was fabricated with a commercial GaAs HBT process. Fig. 5 shows a photograph of the whole transmitter. The size of this chip is 2.6mm x 1.8mm.

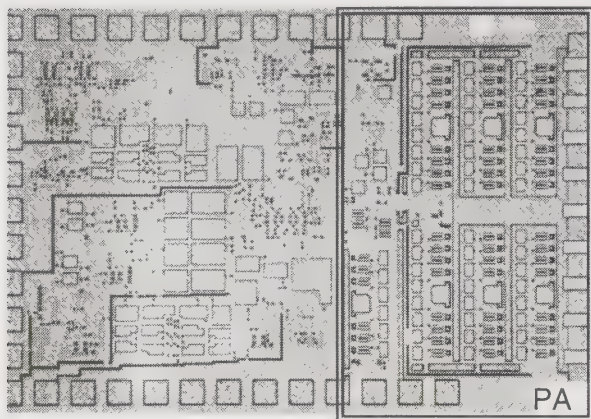


Fig. 5. Chip photograph of a fabricated transmitter. Power amplifier is on the right

Measurements

Because of the long processing time of the ceramic substrate and uncertainties in the device models, we mounted the chips on the regular PCB to check the functional performance of the circuit. The measurements for the amplifier showed that bias circuits worked as expected and an output power of over 2 W was achieved with a supply voltage of 3.5 volts (Fig. 6). The minimum controllable output power achieved at a control voltage of 0.2 volts was -16 dBm. The maximum value of the gain control slope was below 22 V/V, which is in good agreement to simulations. Also an isolation of 40 dB achieved at the nominal 0-dBm input power fulfils our requirements.

The linear gain of the amplifier was 38 dB with a control voltage of 2 volts. However, the measurements for the output stage only indicated, that there are extra losses in the last stage transistor. This explains why the measured efficiency of the power amplifier, 40%, was lower than the value given by the simulations. However, we expect better performance on the ceramic carrier, which is the environment, where the chip was designed to.

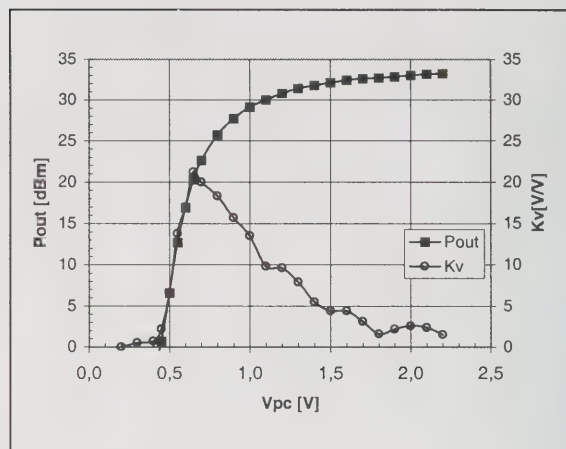


Fig. 6. Measured output power and gain control slope as a function of control voltage V_{pc} at a frequency of 900 MHz ($P_{in} = 0$ dBm, $V_{cc} = 3.5$ V, $V_{reg} = 2.8$ V).

Conclusions

By utilising the regulated voltage available in the mobile phone we get an extra freedom for controlling the output power of the power amplifier. By combining the actual power control voltage with the regulated voltage we can simultaneously realise an effective power switch-off operation and a smooth gain control without compromises.

Acknowledgement

The author would like to thank P. Alinikula, S. Kalajo, K. Nummala, O. Pöllänen, and J. Tanskanen for their contributions in designing the chip and M. Siilin for mounting the chip on the test carrier.

References

- [1] K. Sakuno, et al., "A 3.5W HBT MMIC Power Amplifier Module for Mobile Communications," *IEEE MTT-S Digest*, pp. 63 – 66, 1994.
- [2] W. Abey, et al., "An E-Mode GaAs FET Power Amplifier MMIC for GSM Phones," *IEEE MTT-S Digest*, pp. 1315 - 1318, 1997.

Linearization Performance for a Polar Work Function Predistorter

Claudio G. Rey and Eugene Clark

Spectrian Corporation,

350 West Java Dr., Sunnyvale, CA 94089

408-543-5981, 408-541-0265 (fax), claudio_rey@spectrian.com

Abstract

We present a polar work function predistorter (WFP) which shows good RF amplifier linearization performance and wideband capability. We describe a DSP optimization scheme that allows the linearizer to adapt and maintain optimum out of band performance. We further describe a simulator that accurately models the performance of the system and predicts dynamic distortion due to the WFP implementation as well as an unexpected nonlinear optimum solution for the am-to-am distortion compensation.

Keywords: am-to-am, predistortion, linearization, nonlinear RF amplifier, simulation

Introduction

Correction for RF amplifier nonlinearities goes back to work done on the forties with feedback techniques. Several other techniques have evolved since: one such technique is analog work function predistortion (WFP) which is illustrated in Figure 1. WFP has two distinct advantages over other linearization technologies. One advantage is that the correction is applied before the amplifier where insertion loss can be more easily compensated (unlike feedforward cancellation techniques¹). A second advantage is that the correction architecture is less bandwidth limited than feedback linearization schemes.² The technique can also easily handle a larger bandwidth than current digital predistortion techniques because it is not limited by DSP computational and memory speeds.³

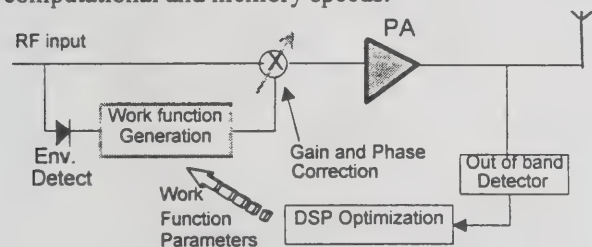


Figure 1. Adaptive work function predistortion.

Predistortion linearizers are designed to produce synthetic (pre)distortion that cancels the distortion generated by the amplifier. RF predistorters achieve such distortion by the use of well-chosen RF device topologies.⁴⁻⁷ WFPs, on the other hand, generate an envelope gain function to correct for the PA nonlinearity--see Figure 1. WFP linearizers are ideal for adaptive compensation of PA distortion with a DSP. These linearizers synthesize either Cartesian^{8,9} or polar¹⁰ work functions and have either been optimized for out of band performance,^{11,8} or for maximum linearity.⁹ The

particular WFP linearizer presented here is a polar work function predistorter that optimizes for out of band performance. The work is based on a patent filed by Spectrian Corporation.

An analog work-function-generation implementation is described as well as a more complete WFP simulator. The WFP simulator includes a spurious detector as well as an optimizing DSP. These last two were also implemented in hardware but are not described here. Results from the simulator demonstrate its importance to understand the optimal performance of the WFP.

Linearizer Equations

Here we present a polar work function linearizer with two work functions. The main advantage of the proposed approach is in its use of a single diode detector and a single squaring circuit. The equation for the gain function generated by the work function implementation is formulated as:

$$G(\rho) = (K + G_1\rho + G_2\rho^2) \exp(j * (L + G_3\rho + G_4\rho^2)) \quad (1)$$

Where 'ρ' is the linear envelope of the input RF signal. The independent terms *L* and *K* are not relevant to the correction scheme. The other four terms are controlled from a DSP optimizer.

Work Function Implementation

A work function predistortion apparatus, as illustrated in Figure 2, was implemented and tested at Spectrian Corporation. The main path for the RF signal is through gain stage A1, the coupler, the delay line (delay1), the vector modulator, gain stage A2, and then on to the PA. The work function generator consists of a linear envelope detector; a low pass filter (LPF), a squaring circuit, four gain-adjust stages (weight adjusters), two summing junctions and finally two high pass filters.

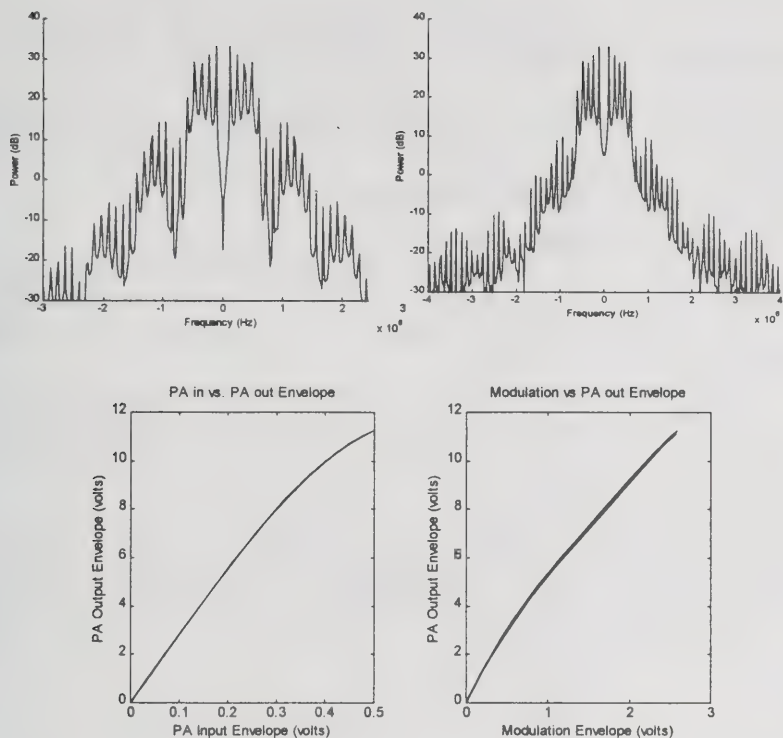


Figure 4. Simulated spectra for a ten-tone signal before (left top) and after (right top) linearization. The ten tones were in phase to generate a signal with exactly 10 dB peak to average.

Figure 5. Simulated (linear) input and output envelope XY plots for the PA (left bottom) and for the corrected PA (right bottom) for the ten-tone signal described above. Note that the plot on the right is 'thicker' due to the filters in the work function. Observe the compression in the PA.

model, (5) an adjacent channel power detector and a triggered DSP subsystem. Both the vector modulator and the amplifier models are derived from experimental data.

The DSP block is triggered to generate perturbations and subsequently to analyze the adjacent channel power resulting from each perturbation. The DSP algorithm requires an insignificant amount of processing overhead. A power detector and a frequency vector analyzer were also included to monitor performance.

Simulation/Lab results

Figures 4 and 5 illustrate ten-tone performance of the WFP linearizer simulator with one of Spectrian's power amplifier behavioural models. The tones were equally spaced and were phase aligned to produce a peak to average of 10 dB. The linearization correction was approximately 7 dB maximum. The power output for

the modeled PA is 44 dBm. At this level the compression level of the amplifier is high due to the peak to average of the ten-tone signal which would require a peak of 54 dBm. Since the amplifier saturates at approximately this power level, the linearization algorithm achieving perfect linearity at this power level is virtually impossible.

Figure 5 illustrates the resulting input/output envelope XY relationship for raw and the corrected PA. These are equivalent to the conventional am-to-am characteristic curves for the uncorrected and corrected amplifiers except that they are obtained as instantaneous envelope XY plots.

The solution provided by the training algorithm is to insert a nonlinearity at lower power levels and to linearize completely at higher power levels--see Figure 5. This results in the envelope XY curve being curved at lower power levels for the corrected amplifier. The envelope XY curve for the corrected amplifier was perfectly linear if the simulation was run for a lower mean output power level. This difference in behaviour at varying power levels demonstrates that the optimal predistorter does not have the inverse am-to-am characteristic of the amplifier when

the amplifier is run near saturation.

The WFP simulator as illustrated in Figure 5 predicts a dynamic effect due to the WFP filters which manifests itself in the "thickness" of the envelope XY plot. The WFP simulator is also a good tool for predicting the peak to average expected at the input stage of the power amplifier. This would not be a simple task otherwise.

Figure 6 a) and b) illustrate the CDMA and two tone performance of the implemented WFP with an RF amplifier driven into strongly biased class AB. Figure 6 c) illustrates a new implementation (different work function formulation not described here) of the WFP hardware with two channels of random IS-95 CDMA data with 13 dB of peak to average, 1.2 MHz bandwidth and 10 MHz separation. As with the ten-tone simulation there was a 7 dB improvement in the third intermodulation product.

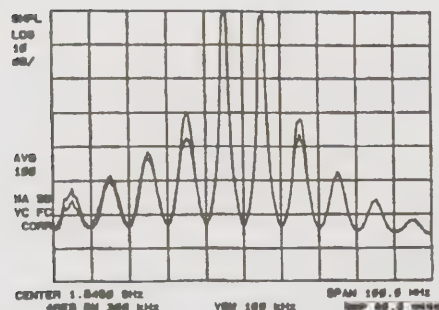
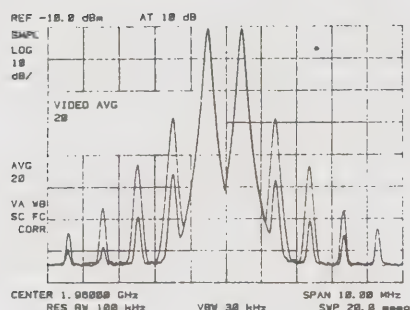
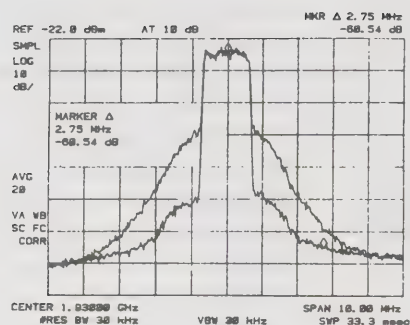


Figure 6. Corrected vs. uncorrected (overlaid) spectral performance a) with a single channel of CDMA and a strongly biased class AB amplifier. b) Same as a) but with two tones. c) Performance of a new implementation with two channels of random CDMA data separated by 10 MHz.

Conclusion

This paper presents a polar work function predistortion (WFP) technique for linearizing RF power amplifiers. The WFP is adaptive and has proven to provide good correction even for a fairly wideband (10 MHz) and high peak to average (13 dB) signal. A WFP simulator is described as well. The WFP simulator demonstrates that the am-to-am of the ideal predistorter is not the

inverse of the am-to-am of the RF amplifier when this is near saturation. The simulator further demonstrates the effect of the WFP filters in the am-to-am characteristic curve for the linearized amplifier. The plots of the implementation spectral regrowth illustrate the potential for the polar WFP to handle both wideband and high peak to average data.

References

- [1] Kenington, P. B.; Beach, M. A.; Bateman, A. and McGreehan, J. P., 1992. "Apparatus and method for reducing distortion in amplification." *US Patent #5,157,345*
- [2] Johansson, M., and Mattsson, T. "Transmitter linearization using Cartesian feedback for linear TDMA modulation", in *Proc. IEEE VTC 91*, pp. 439-444, May 1991.
- [3] Sundstrom, L.; Faulkner, M.; and Johansson, M., "Quantization analysis and design of a digital predistortion linearizer for RF power amplifiers" in *IEEE Transactions on Vehicular Technology*, Vol. 45, Nov. 1996, pp. 707-719.
- [4] Di Benedetto, M. G. and Mandarini, P., "An application of MMSE predistortion to OFDM systems" in *IEEE Trans. on Comm.* Vol. 43, Nov. 1996, pp. 1417-1422.
- [5] Nojima, T. and Konno, T., "Cuber predistortion linearizer for relay equipment in 800 MHz Band land mobile telephone system" in *IEEE Transactions on Vehicular Technology*, Vol. 34, Nov. 1985, pp. 169-177.
- [6] Nannicini, M.; Magni, P. and Oggionni, F., "Temperature controlled circuits for 64 QAM microwave power amplifiers" in *IEEE MTT-S Digest*, 1985, pp. 99-102.
- [7] Namiki, J., "An automatically controlled predistorter for multilevel quadrature amplitude modulation" in *IEEE Transactions on Communications*, Vol. COM-31, May 1983, pp. 707-712.
- [8] Stapleton, S. P. and Costescu, F. C., "An adaptive predistorter for a power amplifier based on adjacent channel emissions" in *IEEE Transactions on Communications*, Vol. 41, Feb. 1992, pp. 49-56.
- [9] Ghaderi, M.; Kumar, S.; Dodds, E. E. "Adaptive predistortion linearizer using polynomial functions" in *IEEE Proc. Communications*, Vol. 141, April 1994, pp. 41-55.
- [10] D'Andrea, A. N.; and Lottici, V. and Reggiannini, R., "RF power amplifier linearization through amplitude and phase predistortion" in *IEEE Transactions on Communications*, Vol. 44, Nov. 1996, pp. 1477-1484.
- [11] Stapleton, S. P. and Cavers, J. K., "A new technique for adaptation of linearizing predistorters," in *IEEE Proc. Communications*, May 1991.
- [12] Simulink is a registered trademark of the Mathworks Inc. of Natick MA.

MMIC-Based Class-A/F Power Amplifier for Ultra-Linear-Commercial PCS Multiple-Modulation-Format Base Station Power Amplification

David Helms, Martina Testa, Nelly Vladimirsky, David Wills

Celwave

4100 Research Way, Corvallis OR 97333

Phone (541) 757-1134; fax (541) 753-8677; email dhelms@celwave.com

Abstract

A MMIC based, three stage, class-A/F driver amplifier provided greatly improved linear power in concurrent multiple modulation formats over previous MMICs and class A/F amplifiers. The modified class A/F structure suppresses the 3rd and 5th order products more than 15.8 dB and 13 dB respectively, over a typical class A/F amplifier, and 17.8 dB over a straight class A amplifier. This provides a 3rd order intercept of 56.7 dBm. For only 48 watts DC. The amp has 0.3 dB gain flatness over a 10% AMPs bandwidth. The technique is patented.

Introduction

As a base station amplifier manufacturer, our AMPS, CDMA, and GSM base stations have typically used Silicon based transistors. Silicon has been preferred due to the BJT's (or LDMOS's) ability to minimize AM/AM distortion, but the BJT's (or LDMOS's) low input impedance forces extensive hand tuning to achieve the very tight gain flatness required by base stations. We modified the currently used class F matching network of a GaAs MESFET MMIC to minimize AM/AM distortion and to take advantage of GaAs FETs higher gain bandwidth product, lower junction temperatures, and the MMICs repeatability.

This modified class A/F amplifier improved linear power over class A and A/F, by using sharp skirted band pass filters coupled with a novel feedback network and very low bias impedances. These changes allowed an significant reduction in AM/AM

distortion. This approach allows the MESFET to be more linear under the high peak to average signals of digital modulation's non constant envelop signals, by reducing high-power shutdown even in multiple modulation formats. The 15 dB peak to average power ratio of IS-97 make CDMA power amplifiers very difficult to design. Si LDMOS or GaAs HBT devices are typically used at 8-9 dB Output Power Back Off due to their high peak current capability. GaAs FETs and BJTs must typically be backed off around 11 dB. The high Output Power Back Offs result in very inefficient power amplifiers. Our design approach allows the GaAs FET amplifier to operate at 6 dB output power backoff.

Analysis

One can mathematically analyze and mathematically eliminate a source of the harmonics and their associated intermodulation products. To estimate the intermodulation products, one calculates the

first to third order derivatives of the tanh function. We found the Taylor series coefficients to be 0 for the DC term, 1 for the f1 term, 0 for the f2 term and -1/3 for the f3 term by solving the equation.

$$y_{(t)} = \{ [A(\sin w_1 t + \sin w_2 t) - 0.33A^3 \{ \frac{1}{2}(1 - \cos 2W_1 t) + \frac{1}{2}(1 - \cos 2W_2 t) + [\cos(w_1 - w_2)t - \cos(w_1 + w_2)t] \} \{ \sin w_1 t + \sin w_2 t \}]^2 \} / 50$$

We find the powers (summed Taylor coefficients) of each carrier to be $(A^2)/50$ and each of the third order products to be $((A^3/4)^2)/50$. While shorting out the harmonic mathematically removes the $2f_2 - f_1$ and the $2f_1 - f_2$ terms. In lab tests loading or shorting the harmonics reduces, but does not eliminate, the intermodulation.

The other case is CDMA modulation. Mathematically, we see that one can analyze the output CDMA signal with respect to the carriers and the harmonics of Multi-tone signals to determine the power coefficients, and take a band pass transform of a signal through³

$$y(t) = F(x(t)) \cos(2\pi f t + \theta) = z(t) \\ (\cos(2\pi f t + \theta) \text{ Where } z(t) = a_1 x(t) + 3/4 a_3 x^3$$

If we use the Wiener_Khintchine theorem we obtain a closed form of the equation⁴

$$P_y(f) = \frac{1}{2} B (P_0 - 6P_0^2 10^{-IP_3/10} + 9P_0^3 10^{-IP_3/5}) + \frac{[3P_0 10^{-IP_3/5}]}{2B^3} (2B^2 - (f - f_0)^2)$$

Where B is the bandwidth, f_0 is the center frequency. The $P_0/2B$ is corresponding to the linear output power density. The rest of the terms are caused by the nonlinearity. These terms cause the ACP. To obtain the IP3 solve:

$$IP_3 = -5 \log \left[\frac{[P_{IM3}(f_1, f_2) B^3]}{[P_0^3 ((3(B - F_1)^3 - (3(B - f_2)^3))] - 4.52 \text{ dBw}} \right]$$

Thus, one can see that CDMA under hard drive has strong 3rd order product

which results in intersymbol interference and the associated BER degradation limiting of the capacity of the PCS CDMA wireless system.⁵

Design

This section shows the design changes to obtain the improved performance over a typical class A/F MMIC.

We had initially designed a class A/F amplifier using traditional design procedures. Its performance was disappointing. It exhibited high drive shut down, poor output power and poor linearity.

To resolve these these problems we manipulated the previously discussed math and found we needed to replace the traditional low pass matching and its harmonic specific poles with sharp skirted band pass matching. We found we improved the class A/F nonlinear product suppression by adding more class F elements, with the most optimal being a five pole band pass filter.

Typical class A/F amplifiers use feedback to provide for oscillation stability. We modified the feedback to go beyond that. The feedback provides instantaneous DC and RF bias control. This feedback improved linearity by reducing the self rectification which shut the transistor under large signal conditions. We found that this was inadequate and needed to further reduce the self biasing by minimizing source inductance. The typical approach of using source wires produced two dB less output power than does the highly via'ed version. It then became clear that we needed to add DC feedback which under large voltage conditions momentarily

upbiases the device to beyond typical class A.

We could upbias the gate but the load line set the maximum current too low to provide adequate peak currents. We found that the load needed to be lower than Cripp's power matching impedance. Furthermore the class F requires the load be predominately negative reactance.

Our class A/F amplifier uses band pass matching on the input and the output as shown in figure 1.

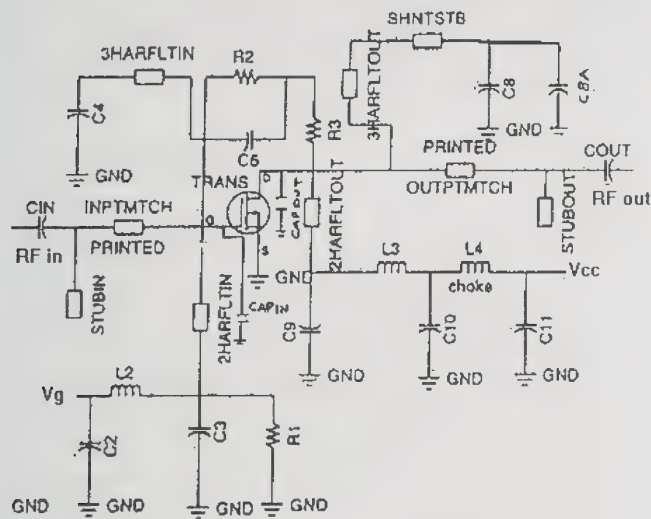


Figure 1. Simplified schematic of class A/F amplifier.

A single stage GaAs MESFET threeway combined class A/F amplifier was modeled, then designed. Simulation and measured data agreed thanks to TriQuint's TOM2 model on Harmonica. To perform the test we operated the amplifier with and without the modified class A/F feedback components. It was power matched, and was operated backed off to a half watt per tone power output. Both amplifiers were operated at the same bias power of 48 watts.

The simple, inexpensive class A/F structure using the feedback linearization suppressed the 3rd and 5th order products more than 15.8 dB and 13 dB respectively beyond that of the class A/F amplifier and 17.8 dB beyond a normal class A 3rd. This is shown in figure 2, It can pass IS_97 at 3W CDMA output power.

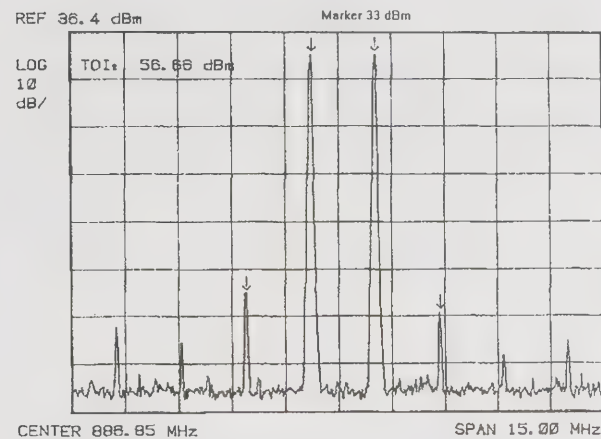


Figure 2. Intermodulation of modified class A/F 12.0 watt base station amplifier.

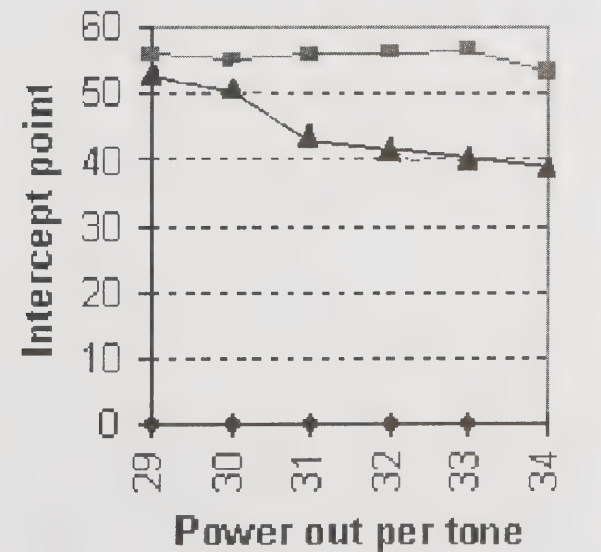


Figure 3. Linearity of Modified Vs non modified class A/F amp Vs power per tone. Figure 3 shows the modified class A/F amplifier (upper trace) demonstrates an increased linearity over the same amplifier

with the class A/F removed (lower trace). Also note, that the linearity degradation knee is three dB higher.

Summary

A highly modified Class A/F amplifier provided dramatically improved linear/efficiency over classes A/F, and A. This was accomplished by reducing harmonics through careful selection of harmonic filters, bias impedance's and multiple feedback networks.

Acknowledgments

Robert Beard, Mehra Mokolla

References

- [1] T.S. Rappaport, Wireless Communications Principles & Practice, Prentice Hall PTR, Upper Saddle River, New Jersey 1996.
- [2] L. W. Couch II, Analog and Digital Communication Systems, Macmillan Pub. Company, New York, 1987.
- [3-5] Qiang Wu, Martina Testa, and Robert Larkin. "Linear RF Power Amplifier Design for CDMA Signals" 1996 IEEE MTT-S Digest.

A High Efficiency MMIC Power Amplifier for Phased Array Antenna Applications

Brian K. Kormanyos*, Ronald W. Kruse, and Debra R. Follensbee

Boeing Information, Space and Defense Systems
Solid State Electronics Development
P.O. Box 3999 MS 3E-36
Seattle, WA 98124-2499

*Phone: (253)773-9247 Fax: (253) 773-9407 Email: brian.k.kormanyos@boeing.com

Abstract

A high efficiency two stage MMIC power amplifier has been designed to meet the requirements of a 64 element transmit phased array antenna operating at 8.2 GHz. The array will provide a high speed data communication link for the NASA GSFC New Millennium program. The array is to be installed on the earth observing scientific satellite EO-1 scheduled for launch in 1999. The two stage MMIC amplifier cell includes an on chip gate bias reference. It provides 23 dB small signal gain, 45% PAE, and requires only 1 square millimeter of die area.

Introduction

The use of phased array antennas for communications and radar applications is expected to increase dramatically in the near future due to advances in MMIC technology and improvements in low cost packaging. Modern phased array antenna systems often include more than 1000 closely spaced modules [1]. The MMIC RF devices for these systems must be small and operate efficiently with a minimum of support components. This helps avoid thermal problems, meet the power constraints of solar powered space and other portable applications, and keeps the modules simple and inexpensive to manufacture.

Circuit Design

A two stage MMIC amplifier has been designed to meet the requirements of a transmit phased array operating at 8.2 GHz. The MMIC is fabricated using the Hewlett Packard 0.25 μm gate length pHEMT process with a 50 GHz f_t and an f_{max} of 100 GHz [2]. Large signal and small signal FET performance is simulated using the Root model [3]. The total gate periphery of the first and second RF stages are 72 μm and 452 μm respectively. The amplifier matching networks are designed to provide a wide small signal and power bandwidth simultaneously. The

amplifier is unconditionally stable and includes on chip power supply bypassing capacitors. Lumped element capacitors and spiral inductors are used to reduce the size of the chip. Full wave electromagnetic simulation of these structures was carried out with Sonnet [4] to insure first pass success.

The amplifier features a simple on chip bias reference. This reference operates from an existing negative 5 Volt supply used by digital ASIC circuitry elsewhere in the module. The bias circuit contains a 10 μm FET and a resistor which act as a self biased current source. The current is forced to flow through a second resistor where the negative gate to source voltage of the small FET is mirrored and used to bias both of the active RF stages. This scheme is largely tolerant of the statistical process variations in the FET pinch off voltage. It is somewhat novel to combine active bias circuitry and RF amplifier stages on the same chip at this frequency range. The power consumption of the reference circuit is low and does not cause a serious reduction in efficiency. The complete two stage amplifier cell including bias circuitry, bonding pads, and on chip power supply bypass capacitors requires only 1 square millimeter of area. A photograph of the complete two stage amplifier cell is shown below in Figure 1.

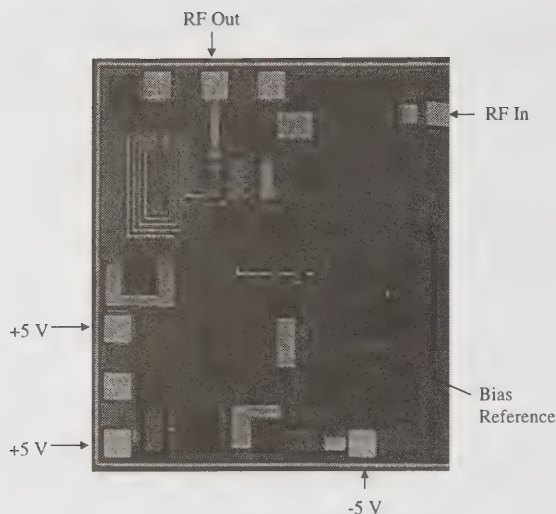


Figure 1: Photograph of the two stage amplifier test cell which is only 1040 x 1040 μm .

RF Performance

The small signal gain of the amplifier exceeds 23 dB with 2dB gain flatness across a 6 to 9 GHz bandwidth. The excellent agreement between

measured and simulated small signal S parameters is shown below in Figure 2. Over a 7.25 to 8.5 GHz power bandwidth the output power at 1 dB gain compression is better than 126 mW (21 dBm) and the saturated output power is better than 200mW (23 dBm). The power added efficiency of the amplifier in saturation is 45 % including the bias reference.

References

- [1] E. Brookner, "Major advances in phased arrays: Part I," *Microwave Journal*, pp. 288-294, May 1997.
- [2] H. Kondoh, et al., "A manufacturable 50 GHz MMIC process and circuit applications," *IEICE Technical Report*, vol.92, No.417, ED92-137, pp.43-48, 1993.
- [3] D. E. Root, S. Fan, J. Meyer, "Technology independent large signal non quasi-static FET models by direct construction from automatically characterized device data," *Proceedings of the 21st European Microwave Conference*, vol. Microwave '91, pp.927-932, Sept. 9-12, 1991.
- [4] Sonnet Software, Inc. 1020 Seventh North Street, Suite 210, Liverpool, NY 13088.

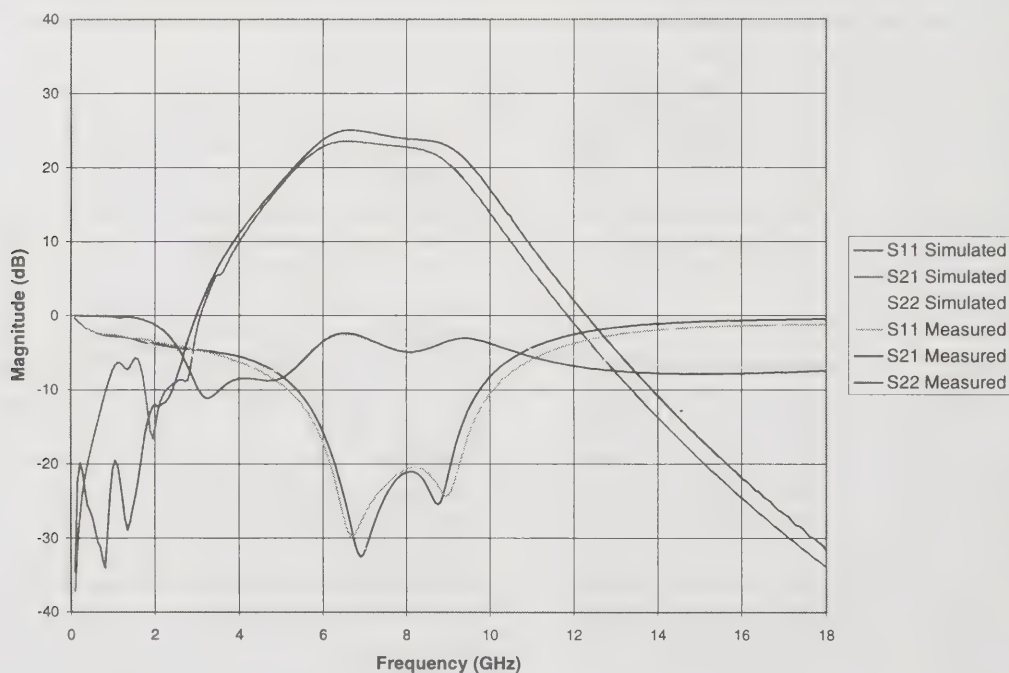


Figure 2 : Small signal S parameter data for the two stage amplifier showing close agreement between simulation and measurement.

Triple Input Single Chip Diversity LNA

Urs Lott

Swiss Federal Institute of Technology (ETH) Zürich,
Laboratory for EM Fields and Microwave Electronics,
Gloriastr. 35, CH-8092 Zürich, Switzerland,
Tel. + 41 1 632 28 11, Fax + 41 1 632 11 98,
e-mail: lott@ifh.ee.ethz.ch

Abstract

For mobile terminals using antenna diversity, a prototype monolithic integrated triple input low noise amplifier (LNA) with diversity function has been realized. The diversity LNA has more than 10 dB gain and less than 4 dB noise figure from 4 to 5 GHz. For comparison, a single LNA was fabricated on the same wafer showing very similar performance. Chip area is reduced by more than 30 % and power consumption by 50 % compared to a multichip LNA with switch.

Keywords: antenna diversity, low noise amplifier, front end, mobile communications

Introduction

There are a number of proposals which require diversity reception in wireless mobile terminals [1,2]. Most of them use separate low noise amplifiers and a diversity switch. Recently, a MCM integration of a diversity front-end was demonstrated [3]. For portable wireless terminals the power consumption and chip size of the receiver front end has to be minimized. In this paper a monolithic one-chip integration of the LNA and the diversity switch function, the "diversity LNA", is presented.

The diversity LNA achieves a minimum noise figure of 3.5 dB at 4.5 GHz with a maximum gain of 14 dB using the enhancement device of a standard E/D-MESFET foundry process. For comparison a single LNA was designed on the same process showing very similar gain and noise performance. Both chips use full on-chip matching circuits. Thus it is shown that the designed diversity LNA requires at least 30 % smaller chip area than a multichip solution without performance penalty.

The diversity LNA uses a single 3.3 V supply voltage for compatibility with low power CMOS circuits.

Power consumption is 13 mW (4 mA from a 3.3 V supply) with one diversity channel

switched on. This is a reduction of roughly 50 % compared to a conventional diversity receiver using separate LNAs.

Circuit Design

The C-band diversity LNA consists of three individual input stages, one for each antenna (see block diagram, Fig. 1).

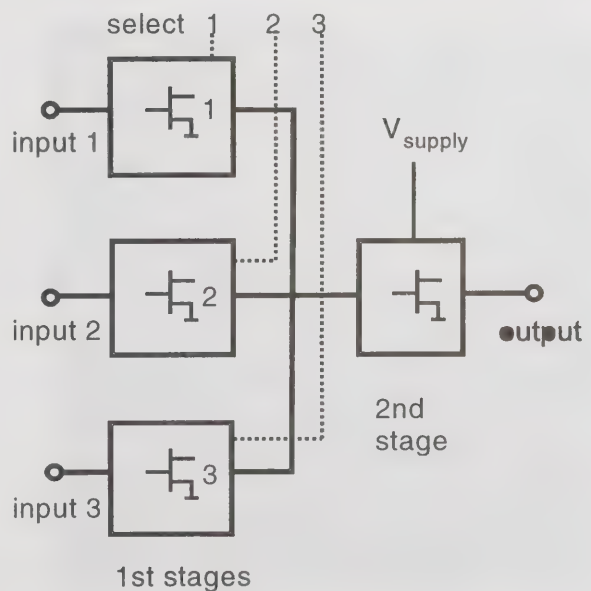


Fig. 1: Block diagram of the three input diversity LNA chip

The bias voltage of the first stage is switched on by the diversity control signal ("select 1,2,3" in Fig. 1). All RF signals are amplified in a common second stage.

The gate DC biasing of the enhancement FET is similar to [4] using very high impedance resistors. Reactive matching is used at the ports and for interstage matching.

Measured Results

For comparison purposes, a single amplifier with a very similar circuit has been fabricated on the same wafer. Fig. 2 shows a microphotograph of the single amplifier having an active chip area of $1.5 \times 0.7 \text{ mm}^2$.

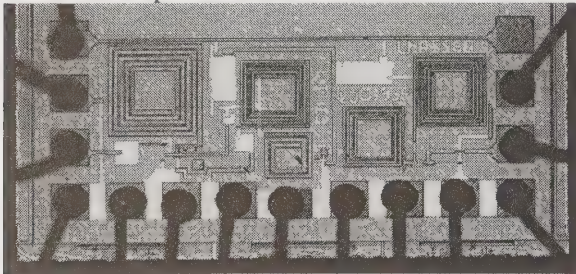


Fig. 2: Microphotograph of single LNA for comparison (size $1.5 \times 0.7 \text{ mm}^2$)

The photo of the diversity LNA is shown in Fig. 3. Its active chip area is about 2 mm^2 , i.e. twice that of a single LNA.

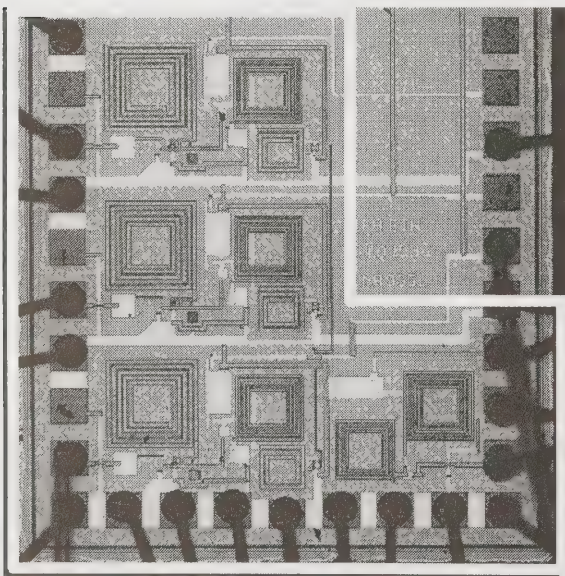


Fig. 3: Photograph of the diversity LNA (size $1.5 \times 1.7 \text{ mm}^2$, active area: 2 mm^2)

The circuits were fabricated on the Triquint QEDA2 foundry process with $0.6 \mu\text{m}$ gate length. Chips from two wafers have been measured.

Fig. 4 shows the gain from input 1 to output when control signal 1 is switched on and off. The gain peak is above 14 dB with an on-off ratio of more than 20 dB, sufficient for many diversity applications.

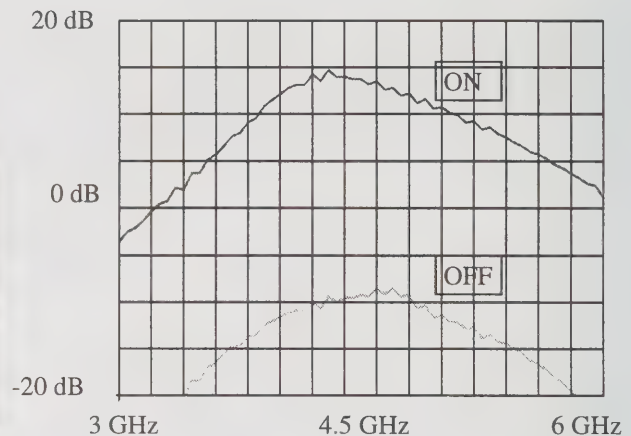


Fig. 4: Measured gain of the diversity LNA from input 1 to output, switched on and off, supply voltage 3.3 V

Fig. 5 shows the noise figure of the three diversity channels activated one at a time. Noise figure (50Ω source impedance) for all three inputs is below 4 dB between 4 and 5 GHz.

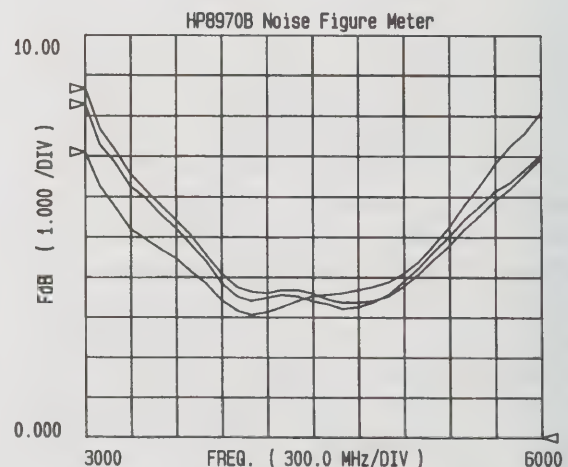


Fig. 5: Measured noise figure of the three inputs of the diversity LNA, switched on one at a time, supply voltage 3.3 V

For comparison, in Fig. 6 the noise figure of the single amplifier, also measured with the HP8970B/HP8971C noise figure meter, is given. The noise figure difference between the single and the diversity LNA is less than 0.2 dB in the range 4...5 GHz, which is within the measurement tolerance. The gain difference between both chips is less than 1 dB (see Fig. 7).

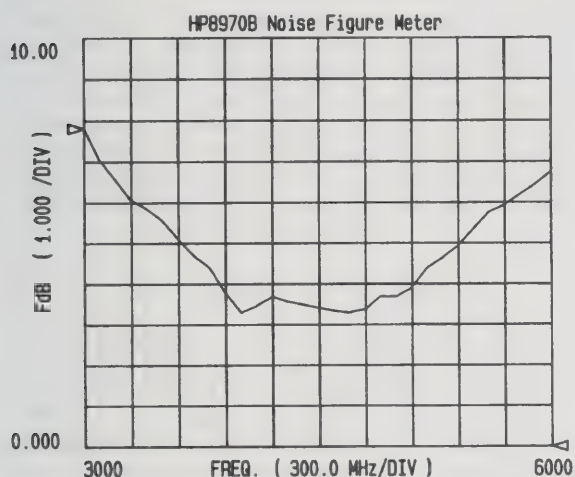


Fig. 6: Measured noise figure of the single LNA at 3.3 V supply voltage

This proves that the diversity function can be added without sacrificing noise or gain performance.

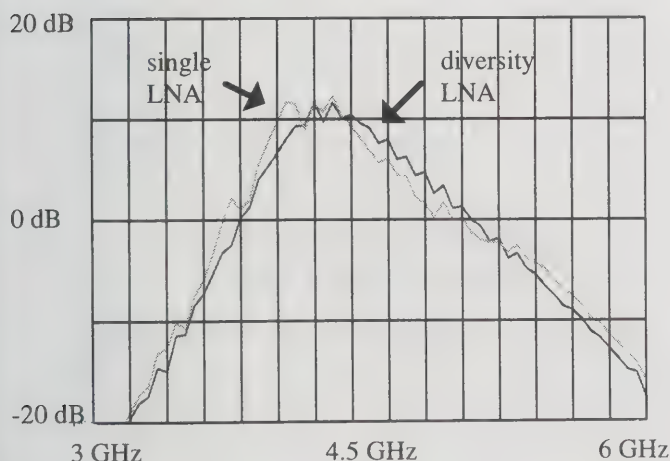


Fig. 7: Comparison of the gain of the three input diversity LNA with the single LNA at 3.3 V supply voltage

Conclusions

A monolithic LNA for diversity applications in wireless mobile terminals at C-band was designed and tested. It has a measured noise figure of 3.5 dB and 14 dB gain at 3.3 V supply voltage, consuming only 4 mA of current with one diversity channel switched on. Compared with a single LNA fabricated on the same wafer, there is negligible performance degradation due to the added diversity switching function. However, chip size of the three input diversity LNA is only about twice that of a single LNA, thus reducing total chip area required for the diversity LNA by at least 30 %. Possible applications for such diversity LNAs include wireless data transmission in indoor and short range outdoor links.

References

- [1] M. LeFevre, M.A. Jensen, M.D. Rice, "Indoor measurements of handset dual-antenna diversity performance," IEEE 47th Vehicular Technology Conf., 1997, Digest vol. 3, pp. 1763 - 1767
- [2] W.-L. Liu, M. Fleischmann, "Advanced low-complexity HIPERLAN receiver using combined antenna switching diversity and simple equalizer," IEEE 47th Vehicular Technology Conf., 1997, vol. 3, pp. 2037 - 2041
- [3] N. Suematsu et al., "An 800 MHz-band front-end MCM for antenna diversity type mobile communication handset terminals," 1997 IEEE MTT-S Symposium on Technologies for Wireless Applications Digest, pp. 167 - 170
- [4] U. Lott, "Low DC power monolithic low noise amplifier for wireless applications at 5 GHz," Proc. of the IEEE 1996 Microwave and mm-wave Monolithic Circuits Symp., pp. 81 - 84

CAD of Microwave Integrated Multistage Active Filters by the Real Frequency Method

P. Jarry, E. Kerhervé, M. Hazouard and J.M. Pham

Microelectronics Laboratory IXL - UMR 5818 CNRS - Bordeaux University

351, cours de la Libération, 33405 Talence - France

Tel. (33) 556 842 806 Fax. (33) 556 371 545

E-mail : jarry@ixl.u-bordeaux.fr

ABSTRACT

This paper presents results of synthesis and optimization for active and integrated microwave filters. The method is based on the real frequency theory. Gain, bandwidth, V.S.W.R. and group delay are the optimized characteristics. Optimization is done with the J.J. MORE subroutine which is an improvement of the Levenberg-Marquardt algorithm. These integrated active low-pass and bandpass filters are used in the field of telecommunications. Applications are given with one or several F.E.T., several equalizers and attenuation poles which are necessary for filters. Lowpass and bandpass examples are given in microwave domains. Responses of filters can have asymmetric frequency characteristics. A CAD program has been written and named FREELCD. We give the different subroutines and show that FREELCD is useful for microwave integrated active filters.

INTRODUCTION

With the techniques of microwave and the growing capacities of satellites, the filters have to be light and less bulky [1]. The presented study is in this area. We therefore give a design of integrated microwave active filters with the aid of Field Effect Transistors (F.E.T.). The described theory is adapted from the « real frequency » method. This technique was introduced for the first time by H. J. CARLIN and B. S. YARMAN [2] and applied in the case of broadband microwave amplifiers. For this theory we only used the measured S_{ij} parameters of each transistor. No equivalent circuit is needed and so it is possible to take into account the non-unilaterality of the used transistors ($S_{12} \neq 0$). In the following section, we show how to transpose the « real frequency » method to synthesize microwave filters.

FORMALISM OF THE REAL FREQUENCY METHOD

A) Synthesis of adaptation ladders [3] [4] [5]

The synthesis is made stage by stage (Figure 1).

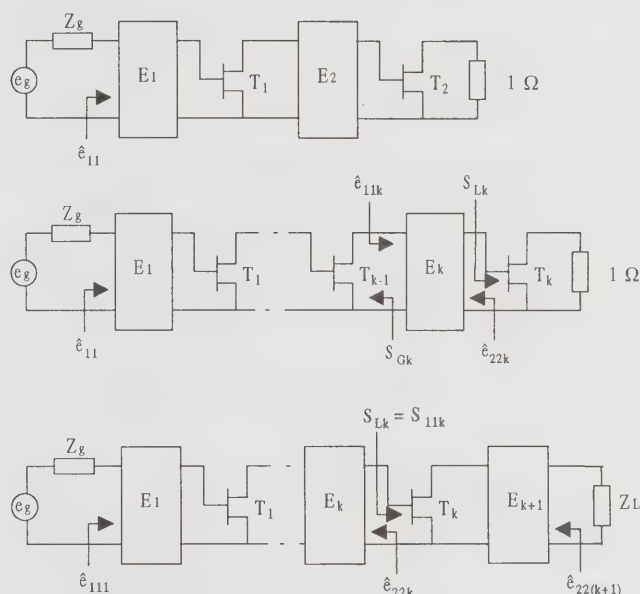
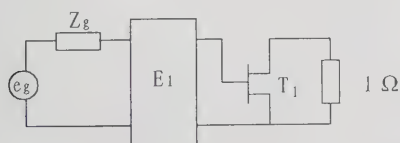


Figure 1. Multistage proceeding.

The first stage is composed of the generator (e_g, Z_g), the equalizer E_1 , the transistor T_1 and 1Ω real normalized load impedance (in general with regard to 50Ω). The equalizer E_1 enables the different optimizations : gain, input VSWR and delay. If these ones are not sufficient, the 1Ω load impedance is replaced by a new block : equalizer E_2 and transistor T_2 loaded by 1Ω . The operation is repeated until we obtain the right results for the gain, input VSWR and phase optimizations. The last equalizer E_{k+1} , loaded by Z_L is placed ; it only optimizes the output VSWR without perturbations on the previous characteristics.

B) Gain, VSWR and group delay

First we measure the scattering parameters $S_{ij k}$ of the FET T_k . The power gain $T_k(\omega)$ of the k first stages depends on the gain $T_{k-1}(\omega)$ of the $(k-1)$ previous stages:

$$T_k(\omega) = T_{k-1}(\omega) \cdot E_k(\omega^2)$$

$$\text{with } E_k(\omega^2) = \frac{|e_{21k}|^2 |S_{21k}|^2}{|1 - S_{Gk} e_{11k}|^2 |1 - \hat{e}_{22k} S_{11k}|^2}$$

e_{ijk} are the scattering parameters of equalizer E_k and S_{Gk} the reflection coefficient at the end of the FET T_{k-1} measured toward the generator.

$$\hat{e}_{22k} = e_{22k} + \frac{e_{21k}^2 S_{Gk}}{1 - e_{11k} S_{Gk}}$$

Overall gain is computed when we have put the last equalizer :

$T(\omega) = (T_1 \cdot T_2 \dots T_k) \cdot E_{k+1}(\omega)$ We apply the same technique to VSWR and group delay.

C) Optimization [6] [7]

Optimization is made simultaneously on the power gain $T(\omega)$, input and output matching $R_{in}(\omega)$ and $R_{out}(\omega)$, and group delay $\tau(\omega)$. We have a least mean squares optimization with the algorithm of J. J. MORE [6]. We have to minimize a multi-objective function:

$$\varepsilon^2 = \sum_{j=1}^N \left[W_1 \left(\frac{T(\omega_j)}{T_0} - 1 \right)^2 + W_2 \left(\frac{R_m(\omega_j)}{R_0} - 1 \right)^2 + W_3 \left(\frac{\tau(\omega_j)}{\tau_0} - 1 \right)^2 \right]$$

T_0 , R_0 and τ_0 are the values of the gain, of the VSWR and the group delay at mid-band. N is the number of discrete frequencies in the bandwidth. W_1 , W_2 , W_3 are the weight functions of the different objectives.

THE FREELCD PROGRAM [8] [9]

The FREELCD program allows us to compute active, integrated filters with the aid of the real frequency method. This method takes into account the non-unilaterality of the FET transistors because we used the measured parameters S_{ij} instead of their equivalent circuit. Stability of the equalizer is always ensured with the Belevitch representation. This method gives us optimization of gain, input VSWR, output VSWR and group delay, as well as the design of multistage devices from narrow band to broad band (Figure 2). Examples are given in the following section.

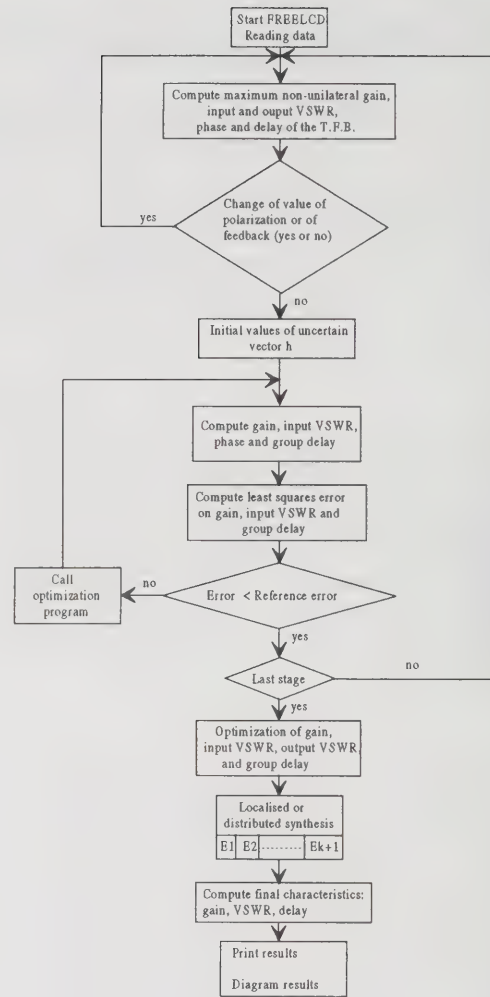


Figure 2. FREELCD configuration.

MICROWAVE EXAMPLES OF INTEGRATED MULTISTAGE ACTIVE FILTERS

All the described proceedings form the program FREELCD [9]. The results are given for normalized frequencies. Initial data (band edges, number of elements, ...) and results are respectively supplied in an input and output file. FREELCD also gives curves of gain, VSWR and delay.

A) First example : low-pass filter [3] [9]

The first application is the synthesis of an active low-pass filter with two stages (two transistors). The measured scattering parameters S_{ij} are given by Alcatel Telecom (France). It is a NEC FET NE20283A, polarized with $V_{DS} = 2V$ and $I_D = 10 \text{ mA}$. This active filter is synthesized in the band $[0 - 4,5] \text{ GHz}$ with two finite attenuation poles ($\omega_0 = 5,8 \text{ GHz}$ and $\omega_0 = 7,5 \text{ GHz}$). So, we have three equalizers (E_1 , E_2 , E_3).

Equalizer E₁: order n=5, without transmission zero

Equalizer E₂: order n=5, with one transmission zero at 7.5 GHz

Equalizer E₃: order n=5, with one transmission zero at 5.8 GHz

After optimization with FREELCD, we obtain the reflection coefficients at the input of each equalizers:

$$e_{111}(p) = \frac{1,46p - 0,86p^2 + 1,84p^3 - 0,39p^4 + 0,45p^5}{1 + 2,76p + 2,74p^2 + 2,52p^3 + 0,88p^4 + 0,45p^5}$$

$$e_{112}(p) = \frac{12,8p - 0,15p^2 + 5,85p^3 + 0,97 \cdot 10^{-2} p^4 + 0,39p^5}{8,33 + 15,75p + 6,05p^2 + 6,14p^3 + 0,47p^4 + 0,39p^5}$$

$$e_{113}(p) = \frac{0,27p - 0,39p^2 - 0,15p^3 - 0,12p^4 - 0,11p^5}{4,98 + 5,82p + 4,39p^2 + 2,14p^3 + 0,69p^4 + 0,11p^5}$$

The gain is still constant in the band of interest; it goes from 19,46 dB to 20,55 dB. This gain is positive in the band-pass [0 - 4,5] GHz while it is negative in the attenuated band (after 4,5 GHz). Input VSWR is less than 1,88 while output VSWR is less than 1,87. Curves of Figure 3 and Figure 4 give gain, VSWR and group delay. Figure 5 gives the final synthesized circuit with two transistors.

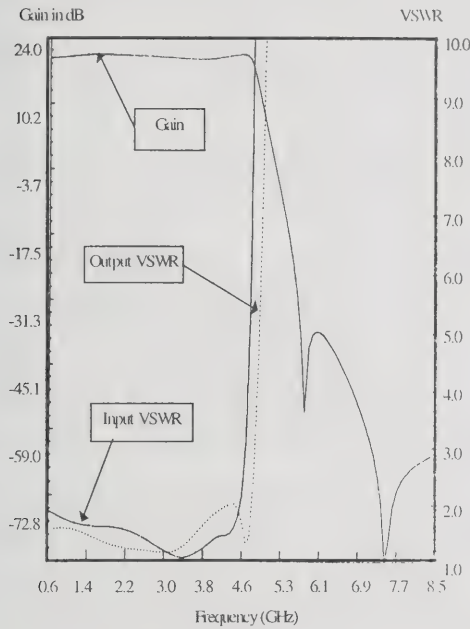


Figure 3. Gain and VSWR of the first example (low-pass)

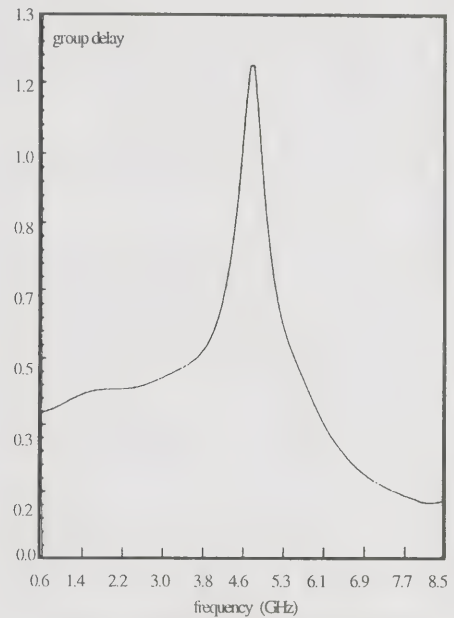


Figure 4. Group delay of the first example (low-pass)

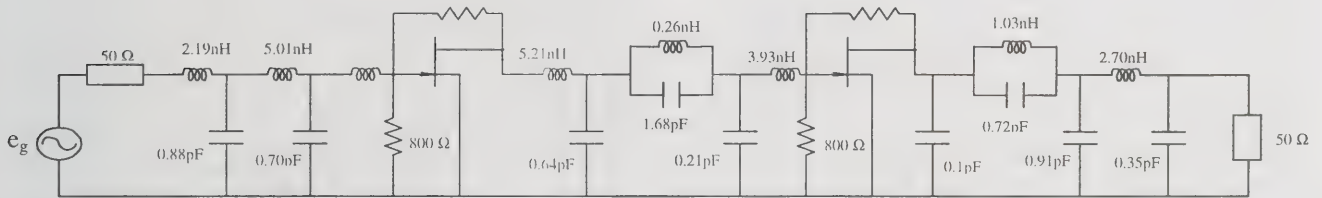


Figure 5. Synthesis of the first microwave integrated active low-pass filter

B) Second example : bandpass filter [3] [9]

With this second example, we have to synthesize an active integrated bandpass filter with only one transistor. The band will be about one octave [2,5 - 5] GHz. There are two finite attenuation poles; one in the left of the bandpass ($\omega_0 = 0,8$ GHz) and one in the right of this same bandpass ($\omega_0 = 6,8$ GHz). We thus obtain two equalizers. After

optimization with FREELCD, input and output equalizers are synthesized from their reflection coefficients $e_{11i}(p)$.

We only give the responses (Figure 6) and the synthesis (Figure 7) of the active filter. The gain goes from 8,26 dB to 8,66 dB in the pass-band. This gain is negative below 2,3 GHz and above 5,35 GHz. The VSWRs are less than 2,0.

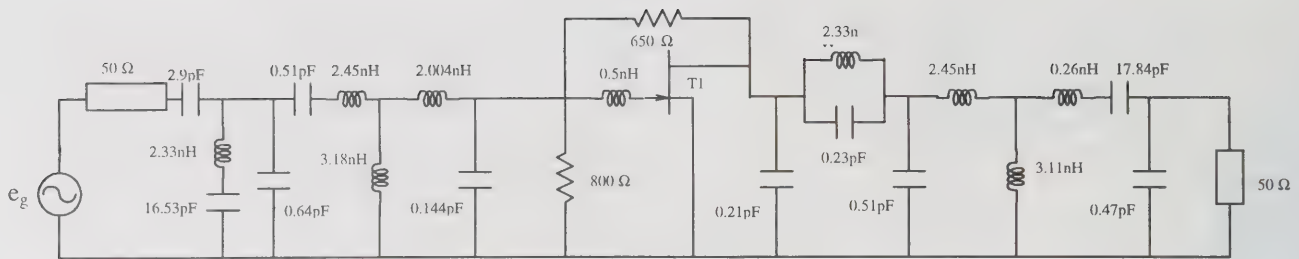


Figure 7. Synthesis of the second microwave integrated active band-pass filter

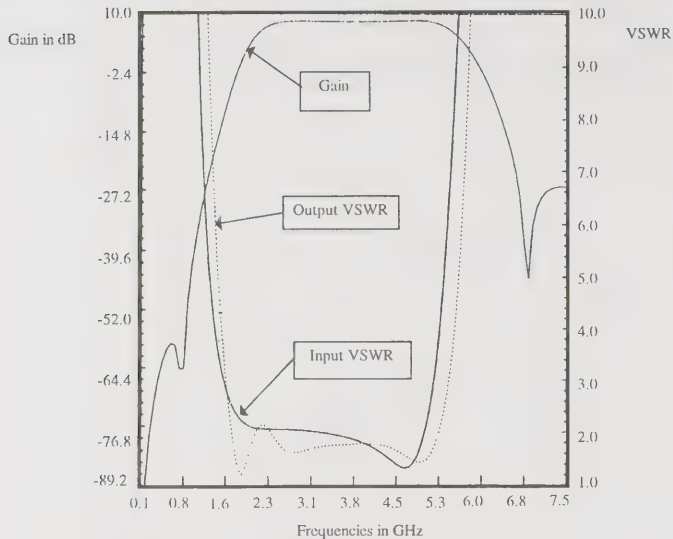


Figure 6. Gain and VSWR of the second example (band-pass)

CONCLUSION

This work is a contribution to the study of integrated active microwave filters. The technique, based on the real frequency method is very strong when it is coupled to a least squares optimization algorithm such as J.J. MORE. From the initial method of H.J. CARLIN, we give an original expansion to broad-band devices when introducing resistive elements of polarization to the transistors. The use of BELEVITCH representation ensures stability. We also take into account the non-unilaterality of transistors ($S_{12} \neq 0$) because the method is based only on the measured scattering parameters S_{ij} of these active elements instead of their equivalents circuits. Consequently this technique enables the optimization of gain, of VSWR (input and output) and of group delay. Synthesis of multistage integrated active filters is now possible from narrow to broad band. We can note that the study of band pass filters is made directly without frequency transformation. The real frequency method (and then FREELCD) can be adapted without difficulties to all kind of transistors

(microwave or low frequencies). « FREELCD » is available.

REFERENCES

- [1] T.T. Ha, *Solid state microwave design*, John Wiley & Sons, 1981
- [2] B. S. Yarman, H. J. Carlin, *A simplified real frequency technique applied to broad-band multistage microwave amplifiers*, IEEE Trans. on MTT, Vol.-30, pp. 2216-2222, Dec. 1982
- [3] E.H. Hendaoui, *Synthèse et optimisation de filtres actifs microondes par la méthode des fréquences réelles simplifiée*, Thèse Télécommunications Hyperfréquences, University of Bordeaux, Oct. 1992
- [4] P. Jarry, A. Perennec, *Optimization of gain and VSWR in multistage microwave amplifiers using the real frequency method*, European conference on Circuit Theory and Design, pp 203-208, ECCTD'87, Paris, 1-4 Sept. 1987
- [5] P. Jarry, *Approximation and synthesis of lumped and distributed filters with arbitrary phase and delay*, IEEE International Symposium on Circuits and Systems ISCAS, New Orleans, USA (invited paper), 1-3 May 1990
- [6] J. J. More, *The Levenberg-Marquardt algorithm: implementation and theory*, Lecture notes in mathematics 630, Springer Verlag, 1978
- [7] P. Jarry, A. Perennec, J. le Bihan, *Optimisation et synthèse d'amplificateurs microondes par la méthode des fréquences réelles*, Revue de Physique Appliquée (Société Française de Physique - CNRS), pp 137-142, Feb. 1988
- [8] E. El Hendaoui, P. Jarry, *The design of microwave integrated multistage active filters*, PIERS, Pasadena, California, USA, July 12-16, 1993
- [9] P. Jarry, E. Kerhervé, H. El Hendaoui, H. Tertuliano, *CAD of microwave integrated active filters by the real frequency - Computer program FREELCD*, Annals of Telecommunications, Vol.51, n°5-6, pp.191-205, 1996

An Examination of Several Large Signal Capacitance Models To Predict GaAs HEMT Linear Power Amplifier Performance

J. Staudinger, M. C de Baca, R. Vaitkus
Motorola, Semiconductor Products Sector
2100 E. Elliot Road, MS EL-712, Tempe, AZ 85284

Abstract

This work examines several methods of modeling charge/capacitance characteristics of GaAs HEMT devices for use in large-signal models applied to the design and simulation of linear power amplifiers. Emphasis is placed on the effect these charge/capacitance models exhibit on the accuracy of the large-signal model to predict linear power amplifier characteristics including gain, output power, efficiency, and linearity. Model predictions are contrasted to on-wafer load pull measurements made using a single-tone sinusoid, two-tone sinusoid, and a $\pi/4$ DQPSK stimuli compliant to the NADC standard. The results suggest the description of the device's charge/capacitance characteristics significantly effects large signal model accuracy.

Key words: GaAs HEMT, large signal modeling of charge/capacitance characteristics

Introduction

Many digital cellular and wireless systems have adopted spectrally efficient modulation formats which require linear, rather than saturated power amplifiers in the transmitter unit of the mobile radio. Performance requirements for these linear amplifiers emphasize spectral linearity while simultaneously exhibiting high efficiency. Linear power amplifiers are inherently less efficient than saturated ones, and thus the linearity-efficiency trade-off becomes a critical design parameter. To improve this design trade-off, efforts have focused on possible improvements offered by higher performance device technologies, such as HEMTs. To this end, accurate large signal models are a critical aspect of the device/circuit design process. This work investigates several methods of describing the charge/capacitance characteristics of a HEMT device for implementation in a large signal model to predict linear power amplifier performance, with emphasis on improving model accuracy. The models are validated by comparing measured large signal load pull data to simulations.

Large Signal HEMT Model

The HEMT large-signal models are developed by considering independently the charge/capacitance-voltage (Q-V and C-V) and current-voltage (I-V) characteristics of the device. A single model is used for the I-V characteristics while three different C-V descriptions are applied to describe charge characteristics. The complete large signal model consists of the I-V expression combined with one of the charge models, resulting in three separate large signal models

I-V Description - HOB D Model This part of the model is based on modifications to a MESFET Curtice model proposed by Halchin *et al.* and is termed the Higher Order Beta Degradation (HOB D) HEMT model [1]. The drain current (I_d) is expressed as a function of intrinsic gate-source (V_{gs}) and drain-source (V_{ds}) voltages, taking the form:

$$I_d = \beta_{eff} [V_{gst}]^{V_{s\exp}} (1 + \lambda V_{ds}) \tanh(\alpha V_{ds}) \quad \text{for } V_{gs} \geq V_t \quad (1)$$

$$I_d = 0 \quad \text{for } V_{gs} \leq V_t \quad (2)$$

where

$$\beta_{eff} = \frac{\beta}{[1 + \mu_{crit} V_{gst}^{g_{m\exp}}]} \quad (3)$$

$$V_{gst} = V_{gs} - V_t \quad (4)$$

and

$$V_t = V_{to} + \gamma V_{ds} \quad (5)$$

The model has eight parameters, α , β , λ , γ , μ_{crit} , V_{to} , V_{gexp} , and g_{mexp} . It is shown in [1] that this model takes into account the degradation in transconductance due to parallel (to the channel) conduction in the low-mobility electron supply layer. Thus, with increasing values of gate-source voltage, transconductance increases until it peaks, and then decreases for larger values of V_{gs} .

Capacitance/Charge Description

The capacitance/charge characteristics are considered independently in three different ways: 1) semi-junction model, 2) modified Statz capacitance model, and 3) table based spline functions.

Diode Semi-junction Capacitance Model: This model is based on diode junction expressions where the device's gate-source and gate-drain capacitance (or charge) are described as functions of intrinsic gate-drain and gate-source voltages, respectively. Note that this model does not include a drain-source voltage dependence. When expressed in terms of capacitance, the model takes the following form:

$$C_{GD,GS} = \frac{C_{GDO,GSO}}{\left(1 - \frac{V_{gd,gs}}{V_{BI}}\right)^{M_{CGD,CGS}}} \quad (6)$$

where C_{GDO} , C_{GSO} , M_{CGD} , and M_{CGS} are model parameters.

Modified Statz Capacitance Model: This model is based on the work of Statz *et al.* [2] with modifications made to partitioning the gate charge. The gate charge is partitioned into gate-source (q_{gs}) and gate-drain (q_{gd}) charge with each exhibiting a dependence on two controlling voltages, i.e., $q_{gs}(V_{gs}, V_{gd})$ and $q_{gd}(V_{gs}, V_{gd})$. Charge is expressed as:

$$q_{gd} = q_1 \cdot (1 - f_P) + q_2 \cdot f_P \quad (7)$$

$$q_{gs} = q_1 \cdot f_P + q_2 \cdot (1 - f_P) \quad (8)$$

The partitioning function f_P takes the form of:

$$f_P = \frac{1}{2} \cdot \left[1 + \tanh(\alpha_Q \cdot v_{ds}) \right] \quad (9)$$

with

$$q_1 = 2 \cdot C_{GSO} \cdot V_{BI} \cdot \left(1 - \sqrt{1 - \frac{V_{NEW}}{V_{BI}}} \right) \quad (10)$$

$$q_2 = C_{GDO} \cdot v_{EFF2} \quad (11)$$

and the remaining terms defined per [2]. Two non-linear transcapacitances are then defined as

$$C_{GDGS} = \frac{\partial q_{gd}}{\partial V_{gs}} \bigg|_{V_{gd}=K}, \quad C_{GSGD} = \frac{\partial q_{gs}}{\partial V_{gd}} \bigg|_{V_{gs}=K} \quad (12)$$

and two ordinary nonlinear capacitances as:

$$C_{GDGD} = \frac{\partial q_{gd}}{\partial V_{gd}} \bigg|_{V_{gs}=K}, \quad C_{GSGS} = \frac{\partial q_{gs}}{\partial V_{gs}} \bigg|_{V_{gd}=K} \quad (13)$$

An equivalent small signal (intrinsic) model illustrating the nonlinear bias dependent elements is shown in Figure 1. Note that transcapacitance is represented as voltage controlled charge sources.

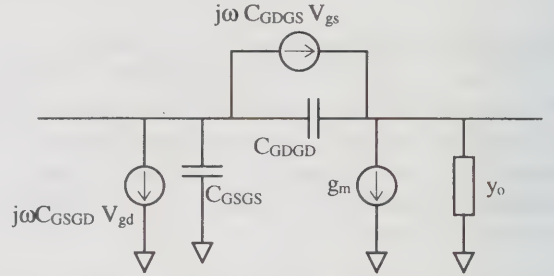


Fig. 1 Small signal intrinsic model illustrating the nonlinear bias dependent elements.

Table Based Spline: This model is based on the work of Root [3] and utilizes spline functions to describe the measured charge characteristics of the device. While it is similar to the previous model (Statz) in that gate charge includes a dependence with drain-source voltage, the mathematical form (splines) is less constraining, thus allowing a possible better representation to measured small signal data. However, spline coefficients must be chosen such that modeled gate charge is path independent/conserving.

Large Signal HEMT Models

Three large signal HEMT models are developed by combining the HOBD I-V description with a capacitance/charge models as illustrated in Table 1. Note that the I-V description and the corresponding model coefficients used to define it are kept identical such that the only difference between the three large signal models is the charge/capacitance description.

Table 1
HEMT Large Signal Models

Large Signal Model	I-V	Capacitance
A	HOBD	Semi-Junction
B	HOBD	Statz
C	HOBD	Spline

Large Signal Model Validation

A GaAs HEMT device was characterized over a wide region of its I-V plane. From this data, large-signal model parameters were extracted for each of the three models. Measurements were made with the device

biased for class A/B operation using on-wafer load pull equipment. Input power was swept over a range to drive the device from linear operation to several dB of gain compression using single-tone sinusoidal, two-tone sinusoidal, and a $\pi/4$ DQPSK NADC compliant stimuli. The source tuner was set to yield good power transfer into the device while the load tuner was set close to a 50 ohm load line (high efficiency for this sized device). Harmonic balance and behavioral analysis simulations were performed using each model. The bias and fundamental/harmonic terminations applied in the load pull system were replicated in the

simulations to accurately describe the measurement conditions. A comparison between measured and simulated characteristics are shown in Figure 2. For a single-tone sinusoidal stimulus, large-signal model predictions of gain and insertion phase based on the semi-junction capacitance model (A) are significantly different than the measured results, especially for input power levels approaching gain compression. Both models B and C more accurately predict gain and insertion phase. Note, however, that under small signal conditions, gain is well predicted with each model. For two-tone sinusoidal and $\pi/4$ DQPSK stimuli, linearity is

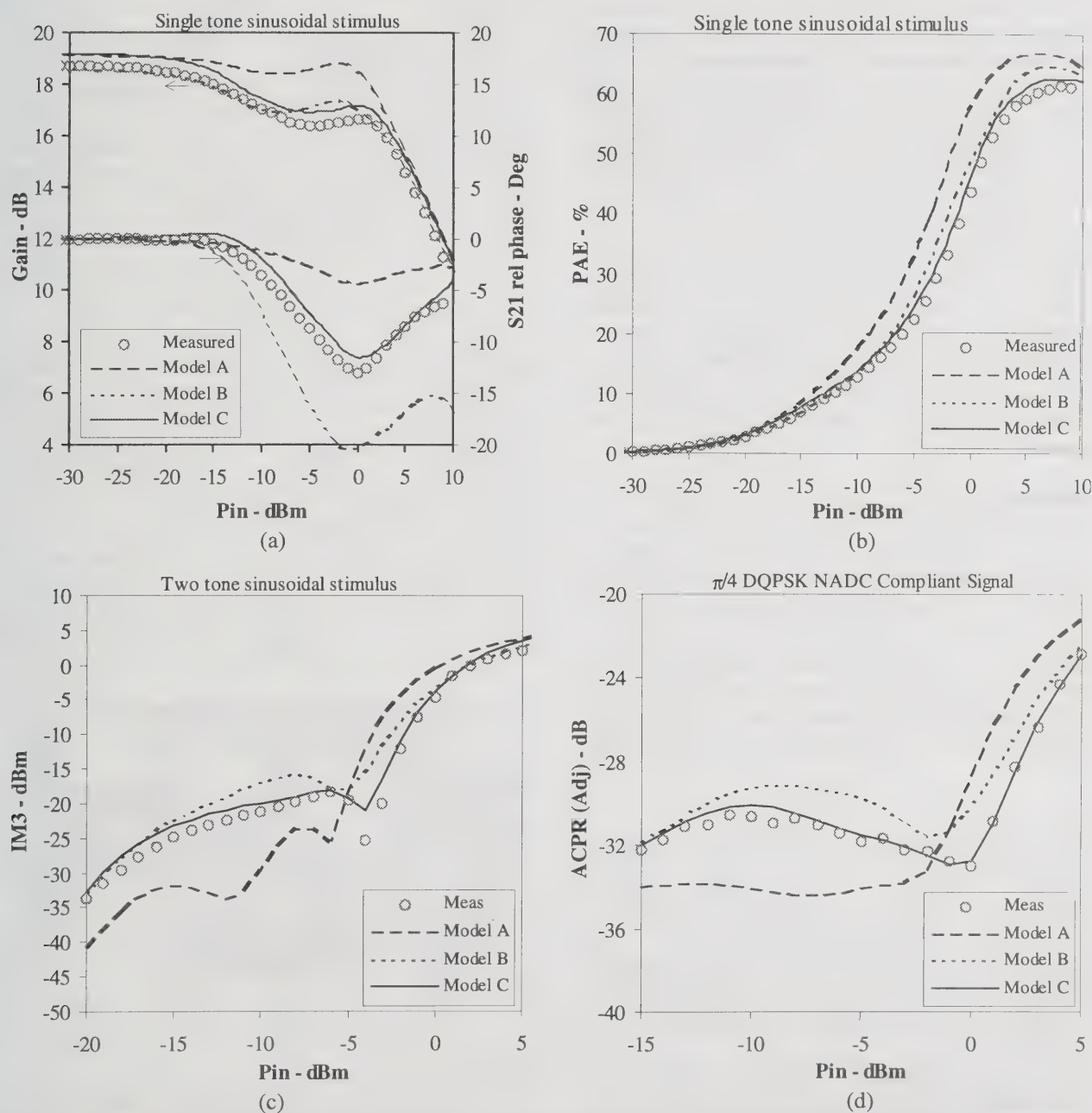


Fig. 2 Comparison of load pull rf measurements to model predictions. a) Gain & insertion phase, b) Power added efficiency, c) Third order IM products, and d) ACPR for an NADC compliant $\pi/4$ DQPSK stimulus.

poorly predicted with model A. An improvement in predictions is obtained with model B while best results are obtained with model C.

The significant differences noted in gain and insertion phase near the area of gain compression between measured data and the three models was investigated further by calculating the dynamic I-V load line. The I-V response is illustrated in Figure 3 for an input power level of -4 dBm. Models A and B predict a response extending over a larger V_{ds} range than model C. This in part explains the higher gain predicted by model A vs model C and measurements. It should be noted that differences in the dynamic I-V load lines are due solely to the charge/capacitance models. In addition to the dynamic I-V response, the dynamic gate-source and gate-drain capacitances were calculated at this load line and input power level (Fig. 3). Note that for model A, these results are calculated with $C_{gd}^* = C_{gd}$ and $C_{gs}^* = C_{gs}$ where as for models B & C, which include transcapacitance terms, $C_{gd}^* = C_{GDGD} + C_{GDGS}$ and $C_{gs}^* = C_{GSGS} + C_{GSGD}$ (reference Fig. 1). It is observed that the $C_{gd}^* - V_{ds}$ and $C_{gs}^* - V_{ds}$ response of models B & C are similar in form where as model A's response is substantially different. This large difference in form is a result of the simplification used in model A where q_{gs} and q_{gd} are functions of single branch voltage V_{gs} and V_{gd} , respectively.

Conclusion

Several methods of modeling charge/capacitance characteristics of a GaAs HEMT device for use in large-signal circuit models has been examined with application to simulating linear power amplifier characteristics. Comparisons between model predictions to rf load pull measurements for several types of stimuli suggest the particular charge/capacitance description of the device has a significant effect on the accuracy of the large signal model.

References

- [1] D. Halchin, M. Miller, M. Golio, S. Tehrani, "HEMT Models For Large Signal Circuit Simulation", IEEE International Microwave Symposium, June, 1994, pp 985-988
- [2] H. Statz, P. Newman, W. Smith, R. Pucel, H. Haus, "GaAs FET Device and Circuit Simulation in SPICE", IEEE Trans. Electron Devices, Feb. 1987, pp 160-169
- [3] D. Root, S. Fan, J. Meyer, "Technology-Independent Large-Signal FET Models: A measurement-Based Approach to Active Device

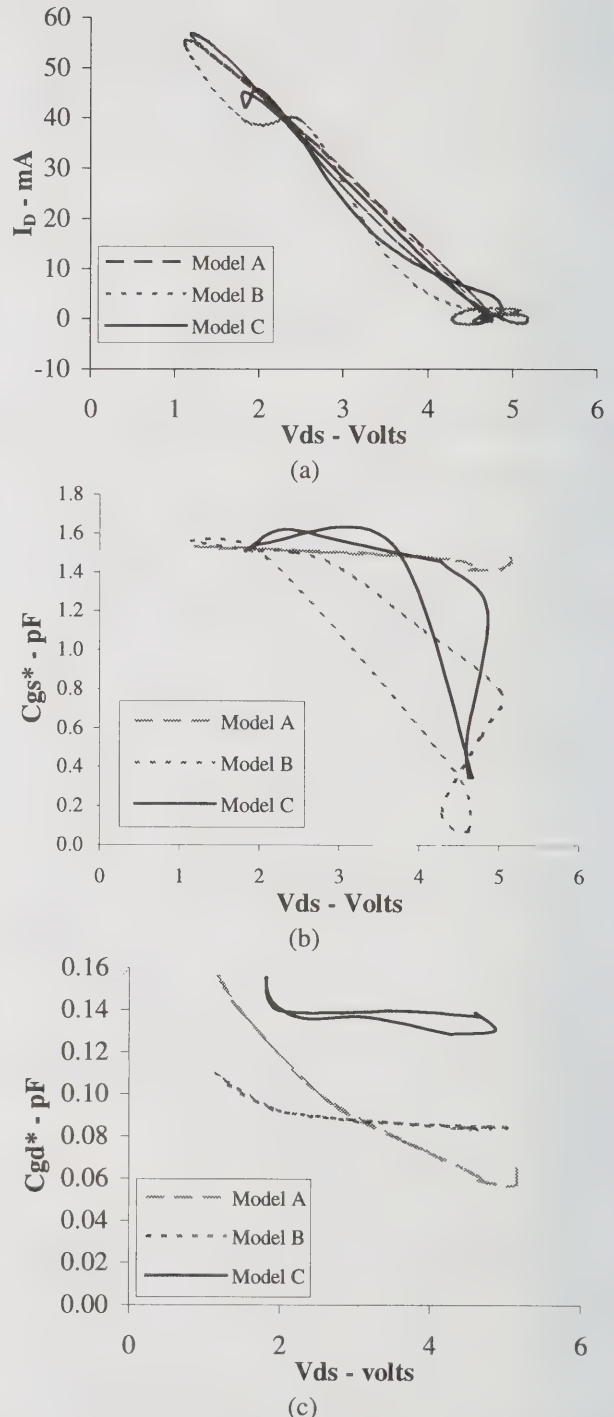


Fig. 3 Calculated dynamic I-V and C-V load lines for an input power level of -4 dBm. a) Dynamic I-V response, b) dynamic C_{gs}^* response, and c) dynamic C_{gd}^* response.

Small-Signal and Noise Modeling of Submicrometer Self Aligned Bipolar Transistor

Salvatore Rinaudo, Giuseppe Privitera, Giuseppe Ferla and Agostino Galluzzo
STMicroelectronics

Stradale Primosole, 50 - 95121 Catania (Italy)

Phone: +39 95 7407585, Fax: +39 95 7407099, e-mail: Salvatore.Rinaudo@ST.com

Abstract—This paper presents an approach to get a full characterization of microwave bipolar transistors starting from S-parameters measurement only. The analytical form of the Y-matrix representation is derived starting from the π -hybrid model [1] and it is used to directly and accurately extract all the parameters of the model. In the second part of the paper the noise sources are added to the equivalent model, and the analytical expression of noise figure is obtained. This expression is used to obtain a full set of noise parameters to characterize the behaviour of the bipolar microwave transistors.

I. INTRODUCTION

RECENTLY, sub-micrometer bipolar transistors for high speed digital and analog circuit applications have been increasing in popularity, and for these devices it is very important to get full small-signal characterization and obtain the device parameters affecting the behaviour of the bipolar circuits. In this paper, the results of the S-Parameters measurement have been used to extract all the information needed for a complete small-signal characterization in terms of gain bandwidth, stability, noise parameters, etc. In the first part of this work the small-signal equivalent circuit was extracted from the Y-matrix obtained from conversion of S-Parameters measurement of the Double Polysilicon Spacer Self-Aligned Transistor (DPSSAT) structure [2] shown in figure 1. In the second part an expression for the noise figure is derived from the small-signal model, and the noise performance are investigated.

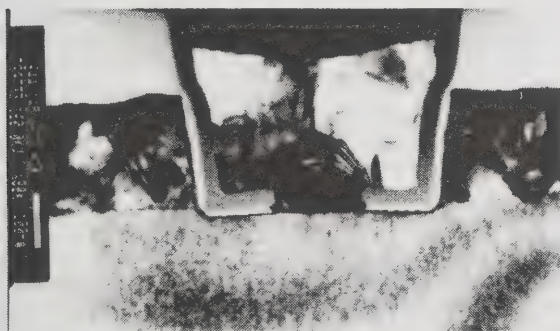


Fig. 1. TEM Cross-Section of the sub-micrometer n-p-n bipolar transistor analyzed

II. SMALL-SIGNAL EQUIVALENT CIRCUIT

To characterize the high-frequency behaviour of bipolar transistors the hybrid- π model [1] (figure 2) is generally used.

Many methods are proposed in literature to extract the element parameters of the equivalent circuit with particular attention to some parameters, such as the base resistance $r_{bb'}$. They generally use a technique based on fitting and optimization of different sets of measurement, such as those in [3], [4], [5], [6], [7]. The technique used in this work is based on a direct method, and all the parameters are extracted at the same time. Considering the circuit of figure 2 as a two-port

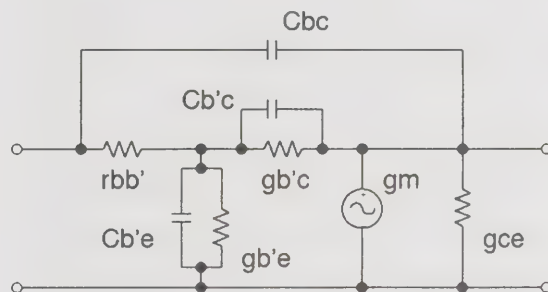


Fig. 2. π -hybrid Small-Signal Model

network, it can be characterized by network parameters. For high frequency (RF region and over) the S-Parameters are often used, because they are easier to measure than the Y or Z or H parameters, but for computer analysis we prefer to use the Y parameters (the other sets can be obtained by conversion formulae). Applying Kirchhoff's laws to the equivalent circuit, and using the emitter as the reference point, the analytical expression of the Y parameters can be obtained in the following definite form

$$\begin{bmatrix} i_b \\ i_c \end{bmatrix} = \begin{bmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{bmatrix} \times \begin{bmatrix} v_{be} \\ v_{ce} \end{bmatrix}$$

where the Y matrix is:

$$\begin{bmatrix} \frac{1}{r_{bb'}}(1 - \frac{1}{r_{bb'}\Delta}) + Y_{bc} & -\frac{Y_{b'c}}{r_{bb'}\Delta} - Y_{bc} \\ \frac{g_{ce} - Y_{b'c}}{r_{bb'}\Delta} - Y_{bc} & g_{ce} + Y_{b'c} + Y_{b'e} \frac{g_m - Y_{b'c}}{\Delta} + Y_{bc} \end{bmatrix} \quad (1)$$

and:

$$\begin{aligned} Y_{b'e} &= g_{b'e} + j\omega C_{b'e} & Y_{b'c} &= g_{b'c} + j\omega C_{b'c} \\ Y_{bc} &= j\omega C_{bc} & \Delta &= \frac{1}{r_{bb'}} + Y_{b'e} + Y_{b'c} \end{aligned}$$

Equating the definite Y-matrix obtained from two-port characterization, in numeric form, to that obtained in analytical form, we have a complex non-linear system of 4 equations (equivalent to 8 real equations) from which the 8 real parameters of the small-signal circuit could be extracted both using a numerical technique such as the Gauss-Newton method [8] or a symbolic solver.

III. SMALL-SIGNAL CHARACTERIZATION

The π -hybrid model is frequency independent and, once it is known at low frequencies, it can be used to analytically simulate the performance of the device up to frequencies of $\frac{f_T}{3}$. For higher frequencies the effects due to distributed phenomena in the base region some errors are introduced. These errors can be seen as an extra phase delay. The bias dependence at $V_{ce} = 0.8$ Volt of some parameters of π -hybrid model is shown. In the figure 3 the base resistance ($r_{bb'}$) is shown as a function of the collector current I_c . To verify the model, in the figures 4 and 5 the microwave behaviour (in terms of Cut-Off frequency (f_T) and S-Parameters) is compared with the experimental data.

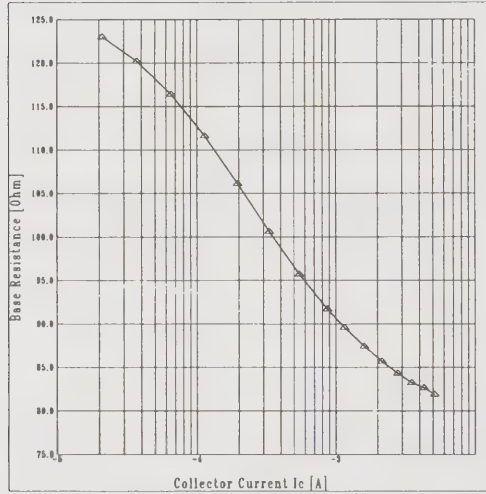


Fig. 3. Base Resistance versus collector current

IV. NOISE MODEL OF THE BIPOLAR TRANSISTORS

Once the π -model is known, the full small-signal equivalent circuit including noise sources can be derived (figure 6). The current generator i_c^2 from the collector to emitter represents a full shot noise due to the random process of the time of arrival of the diffusion and drift carriers at the collector-base junction. Other random processes in the base are the recombination and the carrier injection from the base into the emitter. This is represented by a shot noise current generator; this generator is combined with the Flicker and Burst noise sources, which in bipolar transistors, they can be represented by current generators across the internal base and the emitter, as it has been found experimentally. A thermal noise generator represents the effects of the base resistance, which in the model represents a physical resistor. The collector series resistor shows thermal noise too, but this source can be neglected because it is in series with the high-impedance of the collector node. No thermal noise is associated with $g_{b'e}$, $g_{b'c}$ and g_{ce} because they are fictitious conductances used only for

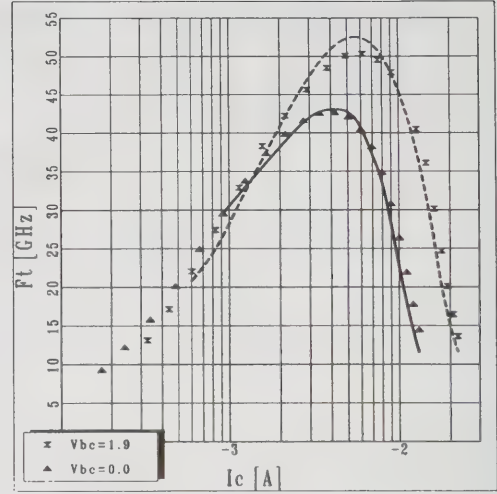


Fig. 4. Gain-Band-Width Frequency Comparison between model data and measurement

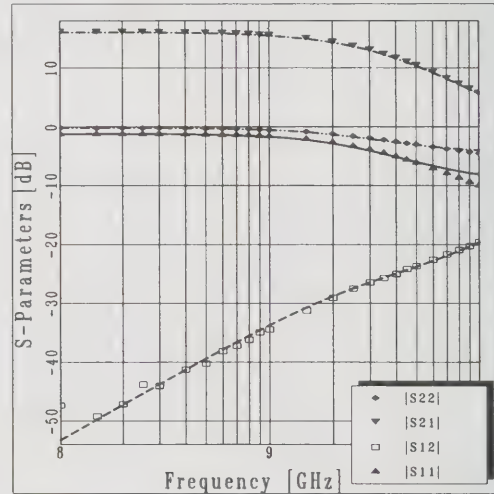


Fig. 5. S-Parameters Comparison at $V_{be}=0.951$ and $V_{ce}=1.0$ Volt between model Data and Measurement, in Bode-Plot

modeling purposes. Then:

$$\begin{aligned} v_b^2 &= 4kTr_{bb'}\Delta f & i_c^2 &= 2qI_c\Delta f \\ i_b^2 &= 2qI_b\Delta f + K_1 \frac{I_b^a}{f^b} \Delta f + K_2 \frac{I_b^c}{1+(\frac{f}{f_c})^2} \Delta f \end{aligned} \quad (2)$$

where k is the Boltzmann's constant, T the temperature, q the electronic charge, f the frequency and Δf is the bandwidth frequency, while the parameters K_1 , a , b , K_2 and c due to Flicker and Burst Noise are constants for a particular device. To evaluate the noise performance, it is necessary to take into account the thermal noise due to the internal conductance (G_s) of the signal source driving of the transistor:

$$i_s^2 = 4kTG_s\Delta f \quad (3)$$

We can consider all these noise generators due to independent factors, so they have no correlation with each other. This

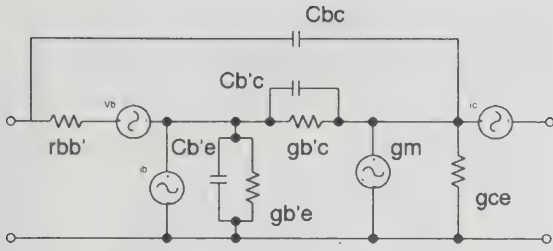


Fig. 6. π -hybrid Full Small-Signal Model with Noise Sources

means that the average value of their product is zero, namely:

$$\overline{v_b \times i_c} = \overline{v_b \times i_b} = \overline{i_b \times i_c} = \overline{v_b \times i_s} = \overline{i_b \times i_s} = \overline{i_c \times i_s} = 0 \quad (4)$$

V. EXPRESSION OF NOISE FIGURE

The definition of the noise figure of a two-port network is well known. It is given as the ratio of the total noise power per unit bandwidth at the output of a system, to the portion of the noise power that is due to the input termination at the standard noise temperature of 290 Kelvin. The same definition can be expressed as of the ratio between the actual total mean squared noise current in the AC-short-circuited output $i_o^2(T)$ and the portion which results from the thermal noise originating in source conductance $i_o^2(S)$.

$$F = \frac{i_o^2(T)}{i_o^2(S)} \quad (5)$$

The total mean-squared current due only to the source conductance is given by:

$$i_o^2(S) = 4kTG_s | \frac{Y_{21}}{Y_{11} + Y_s} |^2 \quad (6)$$

where $Y_s = G_s + jB_s$ is the complex source admittance and the Y_{21} and Y_{11} are component of the complex admittance matrix given in (1). The total mean-squared noise current in the AC-short circuited output $i_o^2(T)$ can be computed by resolving the full-small signal equivalent circuit, including noise sources. This can be done using Kirchhoff's laws and assuming that all the source noise generators are uncorrelated, (equation (4)). Thus

$$i_o^2(T) = i_b^2 \times |A_1|^2 + v_b^2 \times |A_2|^2 + i_c^2 \times |A_3|^2 + i_s^2 \times |A_4|^2 \quad (7)$$

where

$$Y_b = Y_{b'e} + Y_{b'c} \quad Y_x = Y_s + Y_{bc}$$

$$A_0 = r_{bb'} + Y_b^{-1} + Y_x^{-1}$$

$$A_1 = \frac{g_m(r_{bb'} + Y_x^{-1})}{Y_b A_0} \quad A_2 = -\frac{g_m}{Y_b A_0}$$

$$A_3 = -1.0 \quad A_4 = \frac{g_m}{Y_b Y_x A_0}$$

The equation (5) was verified with measurements of noise figure with source impedance of 50 Ω as shown in figure 7.

VI. NOISE PARAMETERS

An alternative expression of the noise figure of two-port amplifiers [9] is given by the formula:

$$F = F_{min} + \frac{R_n}{G_s} |Y_s - Y_o|^2 \quad (8)$$

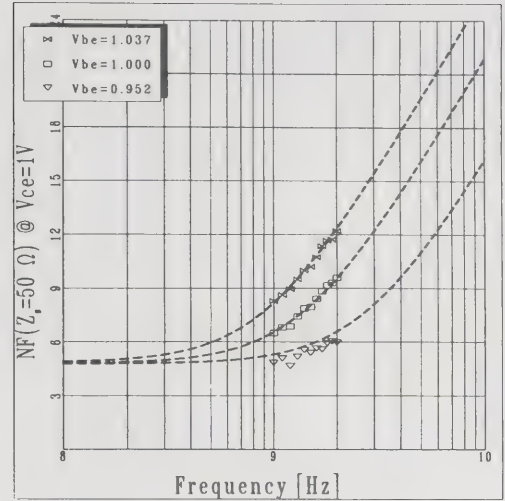


Fig. 7. Noise Figure Comparison with source impedance $Z_s=50 \Omega$ and $V_{ce}=1.0$ Volt

where R_n is the equivalent noise resistance of two-port network, $Y_s = G_s + jB_s$ represents the source admittance and $Y_o = G_o + jB_o$ the optimum source admittance (i.e. where the noise figure has a minimum, called F_{min}).

The previous quantities (F_{min} , R_n and Y_o) are known as **Noise Parameters**. If we know the full set of noise parameters we can completely determine the noise behaviour of two-port amplifiers.

A. Determination of Noise Parameters from AC device simulation

There are many technique to extract the noise parameters described in literature [10], [11]. The method used here starts to compute the noise parameters, after the extraction of the π -model, using one set of formulae already proposed, as those given by Fukui [10]:

$$\begin{aligned} F_{min} &= A + \sqrt{4BC - D^2} & R_n &= B \\ G_o &= \frac{\sqrt{4BC - D^2}}{2B} & B_o &= -\frac{D}{2B} \end{aligned} \quad (9)$$

where

$$A = 1 + \frac{1}{h_{fe}} + r_{bb'} g_m \left[\frac{I_b}{I_c} + \frac{1}{h_{fe}^2} + \left(\frac{f}{f_T} \right)^2 \right]$$

$$B = r_{bb'} \left(1 + \frac{1}{h_{fe}} \right) + \frac{1}{2g_m} \left[1 + (g_m r_{bb'})^2 \left(\frac{I_b}{I_c} + \frac{1}{h_{fe}^2} + \left(\frac{f}{f_T} \right)^2 \right) \right]$$

$$C = \frac{g_m}{2} \left[\frac{I_b}{I_c} + \frac{1}{h_{fe}^2} + \left(\frac{f}{f_T} \right)^2 \right] \quad D = \frac{f}{f_T}$$

where I_b and I_c are the base and collector current, f is the frequency, h_{fe} is the small-signal current gain and f_T is the gain band-width frequency.

The previous formulae are derived from the full small-signal equivalent circuit, neglecting the present capacitances and with further approximations. If we want to determine these parameters in a more accurate way, we can select a range $[Y_1, Y_2]$ of Y_s around Y_o (G_o, B_o) given by (9) and using the equations (5), (6) and (7) to compute the noise figure in n random nodes. These values of noise figure will be fitted to the equation (8), in order to determine more accurate values of the noise parameters.

VII. RESULTS

The proposed method has been used to evaluate the noise behaviour of the sub-micrometer npn bipolar Self-Aligned transistor (DPSSAT). In the figures 8, 9 are plotted the parameters F_{min} , R_n at $V_{ce}=1.0$ Volt, with V_{be} varying from 0.952 V to 1.037 V. While in the figure 10 the constant circles are reported at $V_{ce}=2.5$ V, $V_{be}=0.88$ and frequency equal to 6 GHz.

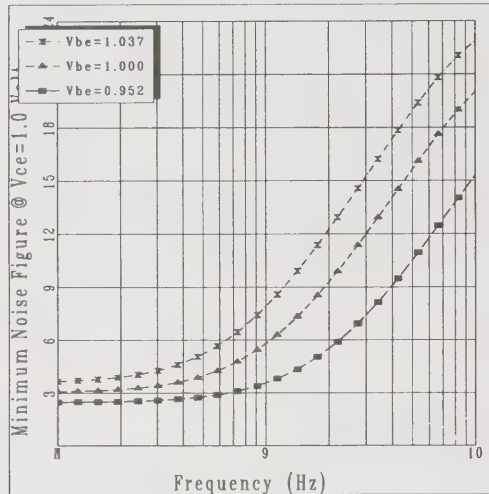


Fig. 8. Minimum Noise figure versus frequency

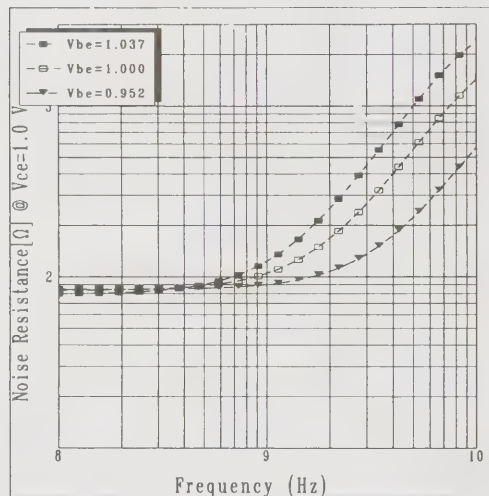


Fig. 9. Noise Resistance versus frequency

VIII. CONCLUSIONS

To complete the information provided by device simulations, small-signal model has been extracted from the Y matrix. This allows us to make a full characterization of microwave transistors by all kinds of network parameters and useful characteristic attributes. The model has been completed by adding sources of noise to obtain a complete characterization in term of NOISE Parameters. The technique and expression proposed have been implemented and verified with

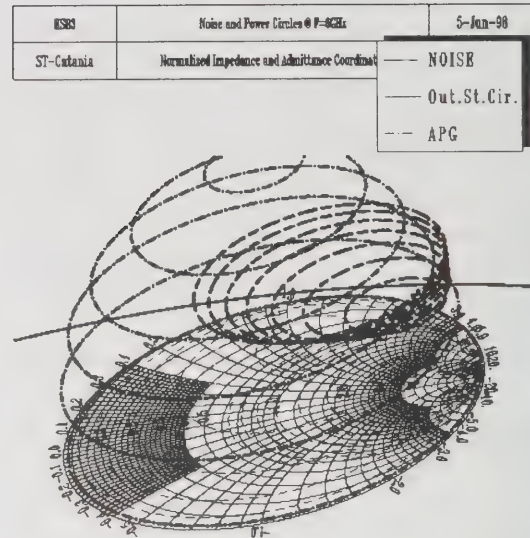


Fig. 10. Power Gain, Stability and Noise constant circles

a sub-micrometer self aligned bipolar device (DPSSAT) using the data obtained from numerical simulation, although the data obtained from experimental measurement could be used as well.

REFERENCES

- [1] L.J. Giacoletto *Study of p-n-p Alloy Junction Transistors from DC through Medium Frequencies* RCA Rev. vol. 15 n. 4 pp. 506-562 December 1954
- [2] Lombardo, S.; Pinto, A.; Raineri, V.; Ward, P.; La Rosa, G.; Privitera, G.; Campisano, S.U. *Si/Ge/SUB X/Si/SUB 1-X/ Heterojunction Bipolar Transistors with the Ge/SUB X/Si/SUB 1-X/ Base Formed* IEEE Electron Device Letters, Vol. 17, Issue: 10, Oct. 1996,
- [3] I. E. Getrew, *Modeling the Bipolar Transistors*, Textronix Inc., Tech. Rep. 1978.
- [4] L. Seonghearn, Byung R. Ryum and Sang Won Kang *A New Parameter Extraction Technique for Small-Signal Equivalent Circuit of Polysilicon Emitter Bipolar Transistors* IEEE Trans. Electron Devices, Vol. 41, NO. 2, pp. 233-238, Feb 1994.
- [5] F. Hebert and D. J. Roulston, *Base Resistance of Bipolar Transistors from Layout Details Including Two Dimensional Effects at Low Currents and Low Frequencies* Solid-State Electronics Vol. 31, No. 2 pp.283-290 (1988)
- [6] F. Hebert and D. J. Roulston, *Computation of the Base Resistance of Bipolar Transistors from Layout Details Including Quasi Three-Dimensional Effects* Solid-State Electronics Vol. 31, No. 8, pp. 1235-1241 (1988)
- [7] T. Fuse and Y. Sasaki *An Analysis of Small-Signal and Large-Signal Base Resistance for Submicrometer BJT's* IEEE Trans. Electron Devices, Vol. 42, NO. 3, pp. 534-539, Mar 1995
- [8] W. H. Press, B. P. Flannery et al. *1986 Numerical Recipes, The art of Scientific Computing* Cambridge University Press.
- [9] H. A. Haus (chairman) *Representation of Noise in Linear Two Ports* IRE Subcommittee 7.9 on Noise, Proceedings of the IEEE, Jan 1960.
- [10] H. Fukui *The Noise Performance of Microwave Transistors* IEEE Trans. Electron Devices, vol ED-13, pp. 329-341, Mar. 1966.
- [11] L. Escotte, J.P. Roux, R. Plana, J. Graffeul and A. Gruhle *Noise Modeling of Microwave Heterojunction Bipolar Transistors* IEEE Trans. Electron Devices, vol. 42, N. 5 May 1995

BSIM3 MOSFET Model Accuracy for RF Circuit Simulation

Suet Fong Tin, Ashraf A. Osman and Kartikeya Mayaram

School of Electrical Engineering and Computer Science

Washington State University, Pullman, WA 99164

Chenming Hu

Department of Electrical Engineering and Computer Sciences

University of California, Berkeley, CA 94720

Abstract

The accuracy of the BSIM3 MOSFET model for small-signal RF circuit simulation has been investigated for a $0.5\mu\text{m}$ CMOS process. Comparisons of the small-signal y and s parameters for different bias conditions and channel lengths show that BSIM3 is reasonably well suited for small-signal analyses up to a frequency of 10 GHz.

1 Introduction

The rapid pace of development in the communication circuits market is driving IC designs to higher levels of integration with circuits that operate at higher frequencies and low voltages. IC designers rely on circuit simulators and accurate models before committing a design to silicon, thus model accuracy is a significant concern for RF applications [1].

The BSIM3 model [2] is an industry standard MOSFET model for deep sub-micron applications. Although the model has been sufficiently validated for DC and low frequency analysis, not much attention has been paid to high frequency modeling. Recent papers [3-6] have started addressing the suitability of BSIM3 for RF applications. In this paper we present a comprehensive study of the accuracy of BSIM3 for small-signal RF circuit simulation in a conventional $0.5\mu\text{m}$ CMOS process. The model is compared with results from two-dimensional (2D) physical device simulation which are accurate for high frequency applications. In this manner, we can evaluate the basic model without accounting for external parasitic elements. Therefore, the substrate network modeling [4-6] is not considered here.

2 Verification Methodology

The methodology used to verify the BSIM3 model is shown in Figure 1. The 2D device simulator MEDICI [7], was used to determine the I - V characteristics, capacitances,

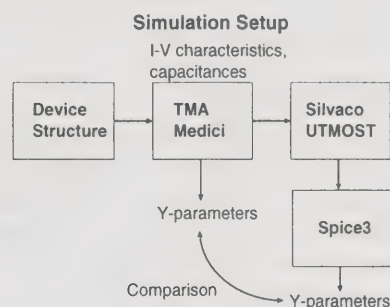


Figure 1: The simulation and verification methodology. The device simulator is used to generate the “measured” data and also for verifying the accuracy of y parameters.

and the y parameters for a MOSFET based on the doping profiles from a $0.5\mu\text{m}$ CMOS process. By using a device simulator, the physics of the model can be readily checked. In this respect, this work is a carefully controlled experiment in which measurement uncertainties have been eliminated. The device simulation results are taken as “measured” data from which the model parameters were extracted using the parameter extractor, UTMOST [8]. Two NMOS transistors of $\frac{W}{L} = \frac{500\mu}{0.6\mu}$ and $\frac{W}{L} = \frac{500\mu}{1.2\mu}$ have been used in this comparison. Both the DC and AC parameters of the BSIM3 model for each device were extracted from the characteristics obtained from the MEDICI device simulator. The fit of the simulated conductance/voltage characteristics from BSIM3 to the device simulator data are shown in Figure 2. The discrepancies in this data reflect as inaccuracies in the low frequency comparisons. It should be noted that the fit in g_{ds} is poor at low gate-source voltages. Once the BSIM3 parameters were extracted, SPICE3 simulations were used to determine the y parameters at different biases, using different options for high frequency modeling. The y parameters were also computed using the device simulator and compared with the y parameters obtained from the circuit simulations.

At high frequencies the nonquasistatic (NQS) operation of the device becomes important [9, 10]. In BSIM3 this effect is accounted for by a first-order model based on the Elmore delay [11]. When set to one, the model parameter NQSMOD introduces the nonquasistatic effect and the ELM parameter controls the value of the Elmore delay. Alternatively, the same behavior can be introduced by use of an external resistance, R_g , in series with the gate terminal as in [3, 5] or by the use of a lumped-sectional model [12]. For this reason, the following cases have been considered in our comparisons: (i) NQSMOD = 0, (ii) NQSMOD = 1 and ELM adjusted to give the correct phase for y_{11} , (iii) external gate resistance, R_g , with NQSMOD = 0 and the value of R_g determined from the phase of y_{11} , (iv) a two-section lumped model with NQSMOD = 0, and (v) the “measured” device data from MEDICI.

3 Results and Discussion

As mentioned in the previous section, the BSIM3 model was verified by comparing the y parameters extracted from the device and the circuit-level simulations. This comparison was made for different biases up to a frequency of 10GHz. A multidevice (lumped-section) approach was used only for the $1.2\mu\text{m}$ transistor, whereby two $0.6\mu\text{m}$ transistors were connected in series to model the distributed channel. The parameters ELM and R_g were found to be dependent on both the channel length and the bias as summarized in Table 1.

Parameter	L=0.6 μm		L=1.2 μm	
V_{gs}	0.8V	3.0V	0.8V	3.0V
R_g (Ω)	0.94	0.7	2.5	1.1
ELM	12.5	1.8	26	4.5

Table 1: Channel length and gate bias dependence of parameters R_g and ELM for $V_{ds} = 3.0\text{V}$.

The y parameters for the $0.6\mu\text{m}$ device for two bias conditions are shown in Figures 3-10. For the $1.2\mu\text{m}$ device y_{11} and y_{12} are shown in Figures 11 and 12, respectively. From the simulated results it is seen that for all five cases a very good agreement is obtained in the magnitudes of the y parameters, whereas there is a disagreement in the phase for some of the cases described above. The following conclusions can be derived from these figures. The parameter NQSMOD affects only the phase of y_{11} and use of NQSMOD = 0 gives an incorrect phase. Cases (iii) and (iv) for incorporating the high frequency rolloff produce results similar to those for NQSMOD = 1. However, NQSMOD = 1 does not accurately predict the rolloff in y_{12} phase, whereas this behavior is captured by both the gate resistance and the lumped-sectional model.

The y parameters are transformed to s parameters and the results from device simulation and BSIM3 with

NQSMOD = 0/1 are compared in Figure 13. Clearly, the agreement in s parameters improves for higher gate biases. From this figure it is again seen that NQSMOD = 1 gives a better agreement between the model and the “measured” data. Since the substrate network has not been included, s_{22} becomes capacitive at high frequencies and the trend is different from those of actual devices [4].

4 Conclusions

The accuracy of the present version of the BSIM3 MOSFET model for small-signal RF circuit simulation has been examined. A good match to physical data can be obtained by use of the NQSMOD flag in BSIM3 and a proper value of the ELM parameter. The same effect can be reproduced by use of an external resistance in series with the gate, or a two-section lumped model can also be used. Since both ELM and R_g are bias dependent, the model is valid only for small-signal conditions. Comparisons of the small-signal parameters show that the core BSIM3 model is accurate up to a frequency range of 10GHz.

Acknowledgements

The authors thank TMA for providing the device simulation tools. This research is supported in part by funding from the NSF Center for the Design of Analog/Digital Integrated Circuits and NSF grant MIP-9702292. The UTMOST parameter extractor was made available through NSF grant DUE-651416.

References

- [1] B. A. A. Antao, “Trends in CAD of analog ICs,” *IEEE Circuits & Devices*, pp. 31-41, Sept. 1996.
- [2] *BSIM3 Version 3.0 Manual*, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, 1996.
- [3] M. C. Ho, K. Green, R. Culbertson, J. Y. Yang, D. Ladwig, and P. Ehnis, “A physical large signal Si MOSFET model for RF circuit design,” *1997 IEEE MTT-S Digest*, pp. 391-394, June 1997.
- [4] W. Liu, R. Gharpurey, M. C. Chang, U. Erdogan, R. Aggarwal, and J. P. Mattia, “R.F. MOSFET modeling accounting for distributed substrate and channel resistances with emphasis on the BSIM3v3 SPICE model,” *1997 IEEE IEDM Digest*, pp. 309-312, Dec. 1997.
- [5] J.-J. Ou, X. Jin, I. Ma, C. Hu, and P. Gray, “CMOS RF modeling for GHz communication ICs,” *1998 VLSI Technology Symposium*, June 1998.
- [6] D. R. Pehlke, M. Schroter, A. Burstein, M. Matloubian, M. F. Chang, “High-frequency application of MOS compact models and their development for scalable RF model libraries,” *1998 Proc. CICC*, pp. 219-222, May 1998.
- [7] Technology Modeling Associates, *Medici*, 1997.

- [8] Silvaco International, *Silvaco Toolset*, 1995.
- [9] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1987.
- [10] R. Singh, A. Juge, R. Joly, and G. Morin, "An investigation into the nonquasi-static effects in MOS devices with on-wafer s-parameter techniques," *Proc. IEEE ICMTS*, pp. 21-25, 1993.
- [11] M. Chan, K. Y. Hui, C. Hu, and P. K. Ko, "A robust and physical BSIM3 non-quasi-static transient and AC small-signal model for circuit simulation," *IEEE Trans. Electron Devices*, pp. 834-841, April 1998.
- [12] Y. Tsividis and K. Suyama, "MOSFET modeling for analog circuit CAD: problems and prospects," *IEEE J. Solid-State Circuits*, pp. 210-216, March 1994.

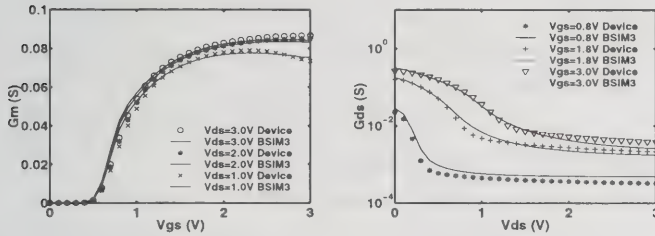


Figure 2: Simulated g_m and g_{ds} from the BSIM3 model and the device simulator for NMOSFET with $\frac{W}{L} = \frac{500\mu}{0.6\mu}$. A reasonable agreement is seen between the "measured" data and the BSIM3 model.

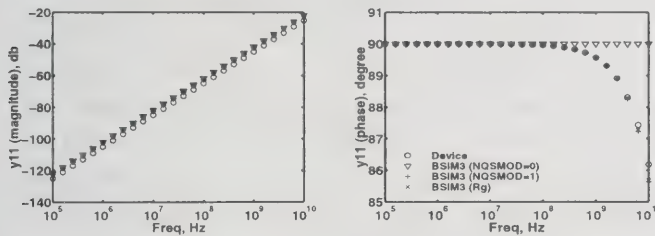


Figure 3: Magnitude and phase of y_{11} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$ with $ELM = 12.5$ and $R_g = 0.94\Omega$. The same ELM and R_g are used in Figures 3-6. The phase roll off in y_{11} is accurately modeled by $NQSMOD = 1$ or by the use of R_g .

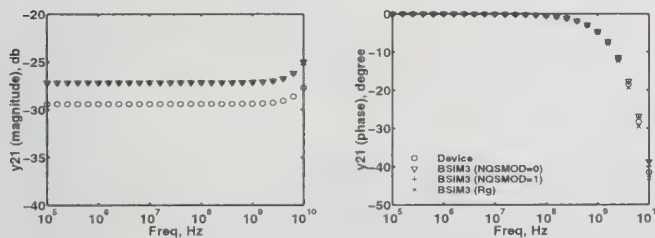


Figure 4: Magnitude and phase of y_{21} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$. y_{21} is accurately modeled even with $NQSMOD = 0$.

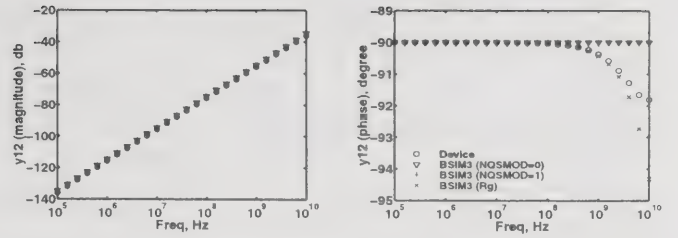


Figure 5: Magnitude and phase of y_{12} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$. The phase roll off in y_{12} is properly modeled only by R_g . $NQSMOD = 1$ has no effect.

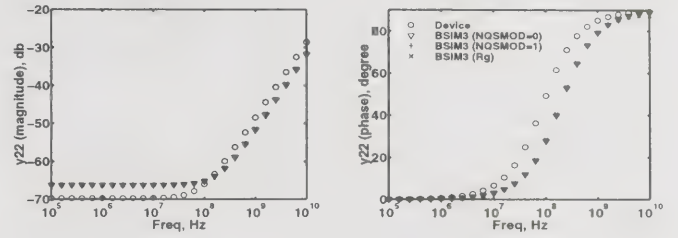


Figure 6: Magnitude and phase of y_{22} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$. y_{22} is accurately modeled even with $NQSMOD = 0$.

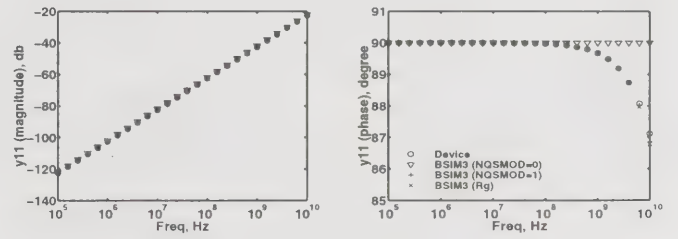


Figure 7: Magnitude and phase of y_{11} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 3.0V$, and $V_{ds} = 3.0V$ with $ELM = 1.8$ and $R_g = 0.7\Omega$. The same ELM and R_g are used in Figures 7-10. The phase roll off in y_{11} is accurately modeled by $NQSMOD = 1$ or by the use of R_g .

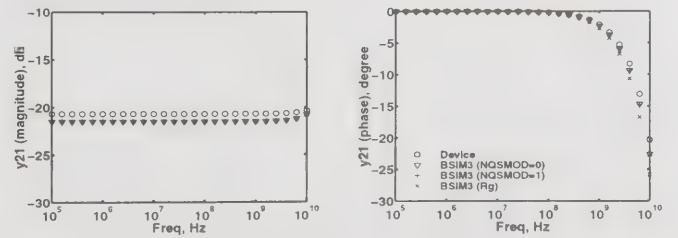


Figure 8: Magnitude and phase of y_{21} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 3.0V$, and $V_{ds} = 3.0V$. y_{21} is accurately modeled even with $NQSMOD = 0$.

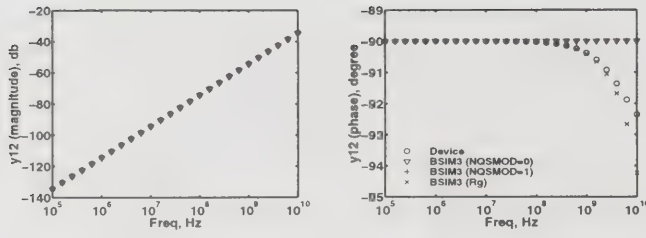


Figure 9: Magnitude and phase of y_{12} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 3.0V$, and $V_{ds} = 3.0V$. The phase roll off in y_{12} is modeled only by R_g . NQSMOD = 1 has no effect.

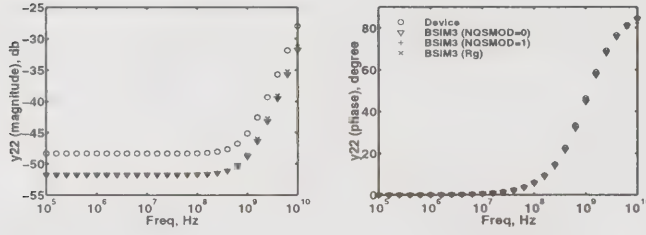


Figure 10: Magnitude and phase of y_{22} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 3.0V$, and $V_{ds} = 3.0V$. y_{22} is accurately modeled even with NQSMOD = 0.

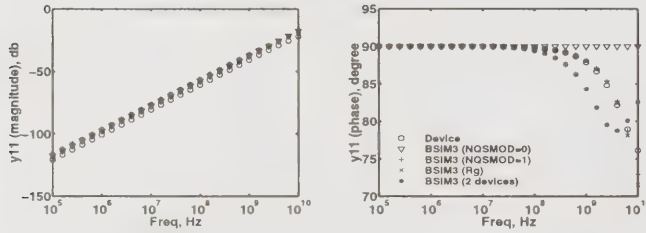


Figure 11: Magnitude and phase of y_{11} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{1.2\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$ with $ELM = 26$ and $R_g = 2.5\Omega$. The same ELM and R_g are used in Figures 11-12. The phase roll off in y_{11} is accurately modeled by NQSMOD = 1, by the use of R_g , and by a two-section lumped model.

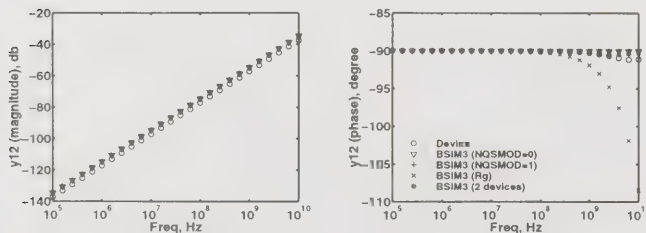


Figure 12: Magnitude and phase of y_{12} as a function of frequency for $\frac{W}{L} = \frac{500\mu}{1.2\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3.0V$. The phase roll off in y_{12} is modeled by the two-section lumped model. R_g overestimates the roll off whereas NQSMOD = 1 has no effect.

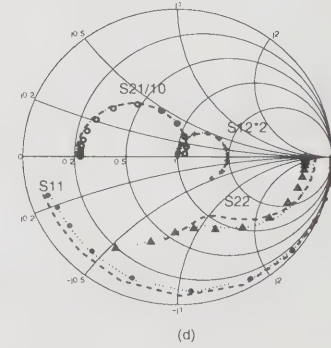
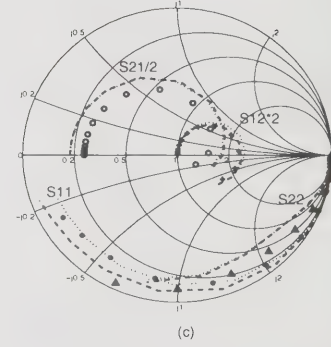
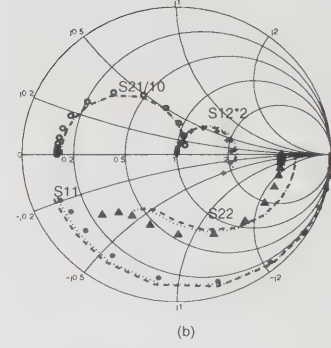
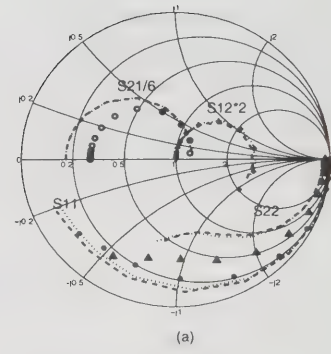


Figure 13: Simulated s parameters from the device simulator (open symbols) and the BSIM3 model: NQSMOD = 0 (dash lines), NQSMOD = 1 (dotted lines) for a frequency range of 100KHz-10GHz. (a) $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3V$, (b) $\frac{W}{L} = \frac{500\mu}{0.6\mu}$, $V_{gs} = 3V$, and $V_{ds} = 3V$, (c) $\frac{W}{L} = \frac{500\mu}{1.2\mu}$, $V_{gs} = 0.8V$, and $V_{ds} = 3V$, (d) $\frac{W}{L} = \frac{500\mu}{1.2\mu}$, $V_{gs} = 3V$, and $V_{ds} = 3V$.

Table-Based FET Model Assembled From Small-Signal Models

C.J. Wei, Y.A. Tkachenko and D. Bartle

Alpha Industries Inc. 20 Sylvan Road, Woburn, MA 01801 Email:cwei@alphaind.com

ABSTRACT

A data-table based large-signal MESFET model is presented based on an ensemble of bias-dependent small-signal equivalent circuits. The model is capable of accurate simulation of small-signal S-parameters over the data-acquisition bias range. In addition to the dc current sources, the model contains two quasi-static charge sources and RF current sources, which are generated by integration of respective intrinsic small-signal element values with respect to the port voltages. All equivalent elements are obtained by cubic spline interpolation. Extrapolation of the model beyond the measurement range is taken into account. The model is extracted by an in-house software without involving optimization. The validity of the model is demonstrated by comparing the simulation of small-signal S-parameters over a wide bias range and power performance to the measured data.

INTRODUCTION

On-growing challenges in wireless communications impose a great pressure to the circuit designers for releasing a new product in a shortest possible cycle time. It is imperative to have a versatile model that can provide accurate prediction of small-signal response as well as power performance. Analytic empirical nonlinear models, which are widely used, require good model formulation and optimization process in extraction to fit the measured data. Table-based approaches have demonstrated their ability to accurately predict device performance and they are technology-independent [1,2,3]. They also eliminate optimization in model extraction process, making thus the model versatile and efficient.

MODEL DESCRIPTION

The new model is derived from the well-accepted small-signal models. A 13 bias-dependent element equivalent circuits are used to represent the small-signal model. The element list includes 6 extrinsic resistances and inductances and 7 intrinsic elements, as shown in figure 1. The extrinsic elements are extracted from the COLDFET approach and the intrinsic elements are directly calculated from measured network parameters at various bias points. The values of bias-dependent small-signal equivalent circuit elements are analytically generated without involving optimization. The extraction is, therefore, very efficient.

In the large-signal equivalent circuit, as

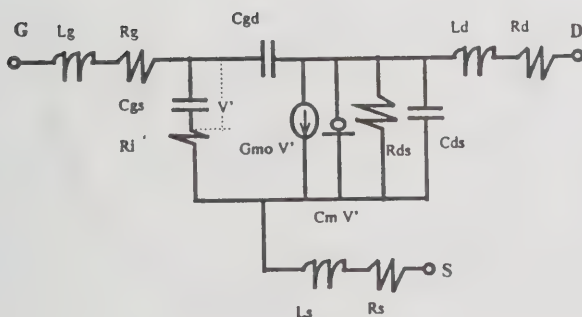


Figure 1. Small-signal Equivalent circuit of MESFETs

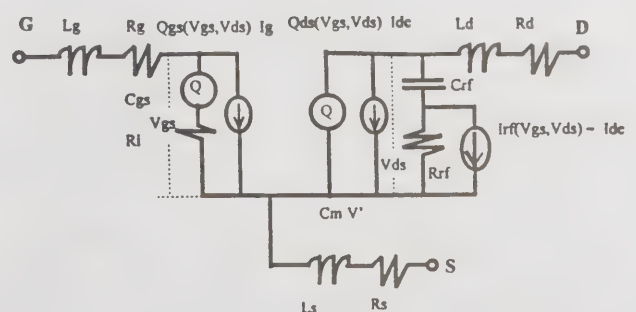


Figure 2. Large-signal Equivalent circuit of MESFETs

shown in figure 2, two charge-sources Q_{gs} , Q_{ds} and two current sources are used. The charge-sources are generated by integration using three capacitance's C_{gs} , C_{gd} , C_{ds} and one trans-capacitance C_m . The integration expressions are:

$$Q_{gs} = \int C_{gs} dV_{gs} + (-C_{gd}) dV_{ds}$$

$$Q_{ds} = \int (C_m - C_{gd}) dV_{gs} + C_{ds} dV_{ds}$$

It has been verified that the integration of charge terms is path-independent.

The RF current characteristics are, however, different from dc curves. The difference comes from low-frequency dispersion effects such as self-heating and deep level trapping. Again, the RF I_{ds} can be obtained by the following integration,

$$I_{dsrf} = \int G_m dV_{gs} + G_{ds} dV_{ds}$$

Since the dispersion effects vary with biasing during S-parameter measurements, the integration may show non-physical path-dependence [4]. However, by choosing a special integration path, this difficulty can be bypassed. The special path assumes that the measured G_m in the linear region and the measured G_{ds} in the saturation region are the real RF response parameters and they are not affected by the change of dispersion while biasing. The RF current source is incorporated into the drain port of the device via a series RC circuit as shown in figure 2, to account for the dispersion time delay.

MODEL EXTRACTION AND IMPLEMENTATION

A set of 450 bias-dependent S-parameters is measured on a 2-mm GaAs MESFET. The bias range is $V_g = -2.2$ to 0.6 V and $V_d = 0$ to 5 V. The measured multiple S-parameter data are fed to our in-house software for model extraction.

A user-defined element was written in the framework of LIBRA for model implementation. The data table is first read and all the element values, including charge and current terms and conductance/capacitance

terms are calculated by cubic spline-interpolation.

The model is extrapolated beyond the measurement region by assuming all conductance and capacitance terms to have the same decreasing factor. For instance, the drain current beyond the measurement range, $V_d > V_{dm}$, is expressed as

$$I_{ds} = I_d(V_{dm}, V_{gs}) (1 + G_{ds}(V_{dm}, V_{gs}) V_t (1 - \exp(-(V_{ds} - V_{ds})/V_t)))$$

where V_t represents the extend of V_{ds} beyond the measured range where G_{ds} approaches to zero. Another parameter, V_1 , is used in extrapolation of the model to the region $V_{gs} < V_p$.

MODEL VALIDATION

The small-signal simulation at 66 bias points over the measurement range ($V_g = -1.4, -1.2, -1.0 \dots 0.6$ and $V_d = 0.5, 1, 2, 3, 4, 5$) is performed and then compared to the measured S-parameters. Figures 3 and 4 show the modeled and measured S-parameters at near-linear region ($V_d = 0.5$ V and $V_g = -1$ V) and saturation region ($V_d = 4$ V and $V_g = -1.2$ V). It is seen that the fitting is very good. Figures 5(a) and 5(b) show the respective discrepancy in S21 (magnitude and phase)

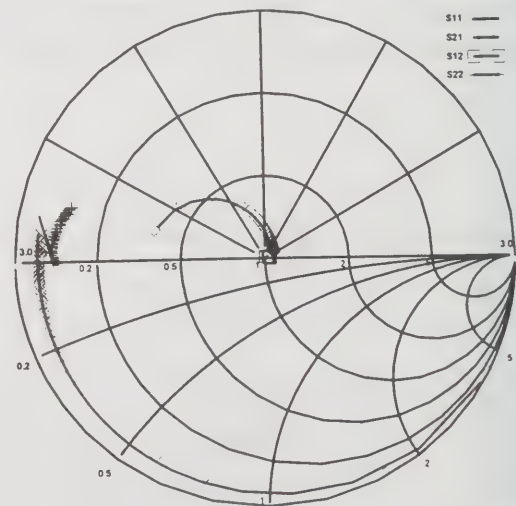


Figure 3. Modeled (lines) vs. Measured (symbols) S-Parameters. $V_{ds} = 0.5$ V and $V_{gs} = -1.0$ V

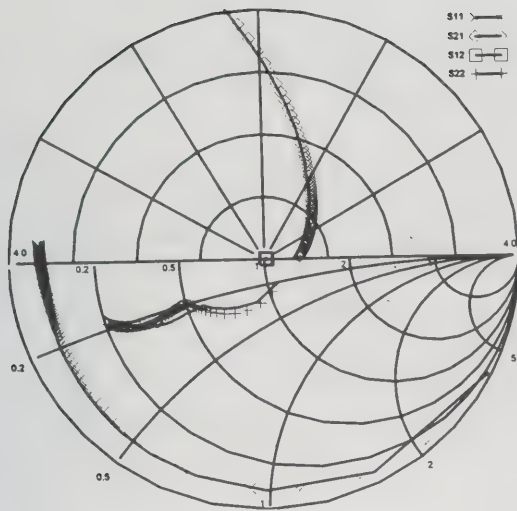
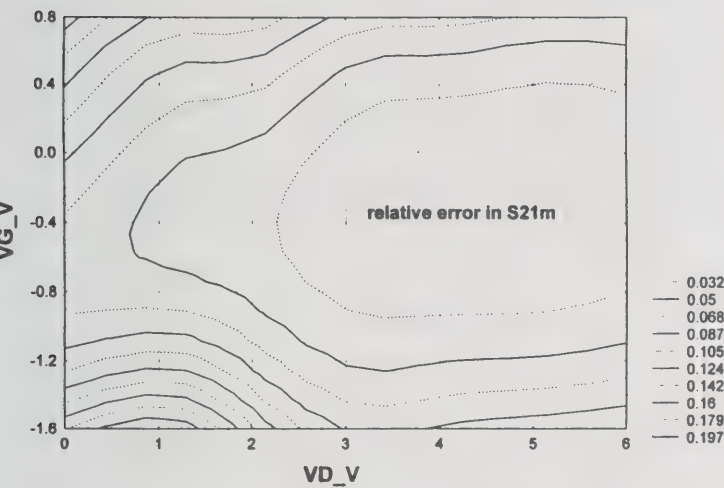
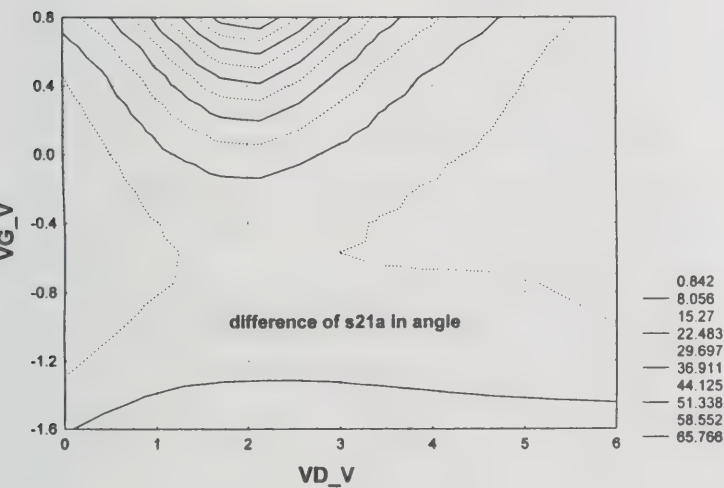


Figure 4. Modeled(lines) vs. Measured(symbols) S-parameters. $V_{ds}=4$ V and $V_{gs}=-1.4$ V.



(a)



(b)

Figure 5. Relative error in $\text{Mag}(S_{21})$ (a) and $\text{Ang}(S_{21})$

at $f=1$ GHz. Similar errors are obtained also for S_{11} , and S_{22} (not shown here). Again, the errors over around 80% area of the measurement range are small. The typical errors are 2.5% for $\text{Mag}(S_{11})$ and 11° in $\text{Ang}(S_{11})$, 7% for $\text{Mag}(S_{21})$ and 8° for $\text{Ang}(S_{21})$, and 5.4% for $\text{Mag}(S_{22})$ and 14° for $\text{Mag}(S_{22})$. Obviously, the error originates from the small-signal model extraction.

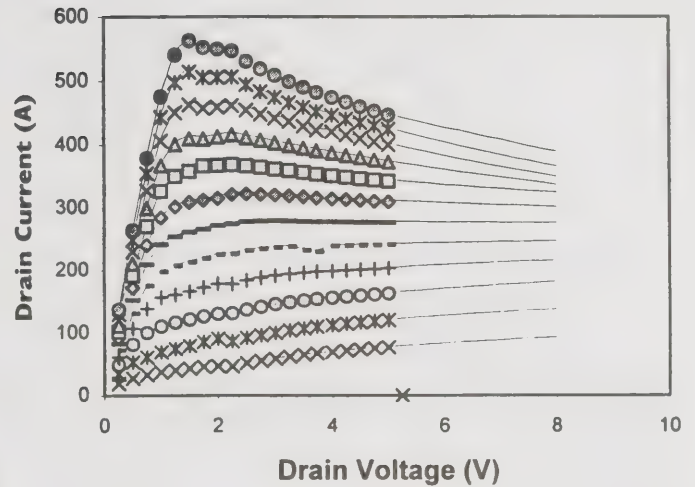


Figure 6. Modeled(lines) vs. Measured (symbols) dc-curves

The dc curves are simulated up to $V_{ds}=8$ V and are shown in figure 6. There is a perfect fitting in the measurement region. Power simulation was carried out at $V_{ds}=3$ V and $V_{gs}=-1.2$ V and at input power up to 12 dBm. The results are shown in figure 7. It has been verified from waveform measurements, that the drain voltage swings beyond the measurement range at $P_{in} \geq 0$ dBm. Therefore, the power performance at higher power is strongly dependent on the extrapolation. Normally it is difficult to have exact measurements at high bias range without risking the damage of device. Therefore, the simulation beyond the data acquisition region is not expected to be perfectly matched the measured data. However, it does give qualitative consistent results, especially near the extrapolated region. At lower power range, the simulation agrees with the measured data very well.

To verify the model at high power region, the measurement region is extended to $V_{ds} = 8V$. Figure 8 is the plot of output power, second and third harmonics and power added efficiency versus input power. The ripples of the

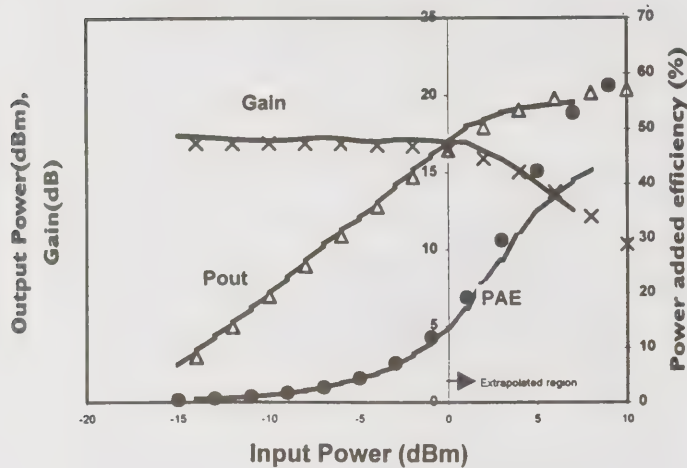


Figure 7. Modeled(line) vs. Measured (symbols) power Performance. Quiescent point: $V_{ds}=3V$ and $V_{gs}=-1.2V$.

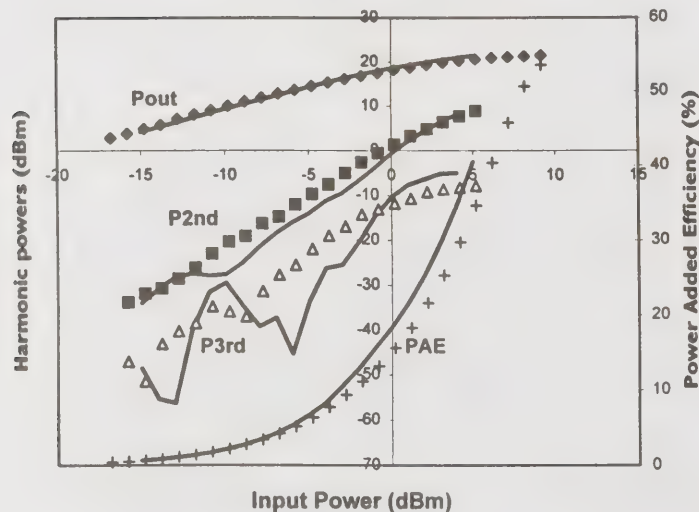


Figure 7. Modeled(line) vs. Measured (symbols) power Performance. Quiescent point: $V_{ds}=3V$ and $V_{gs}=-1V$

third harmonics at lower power are predicted but smaller than the measured. At higher power, somewhat higher than measured PAE and third harmonic level is predicted. The inaccuracy is due to the fact that V_{ds} swings to the

extrapolated region below pinch-off, which was not yet fully characterized.

CONCLUSIONS In this paper we presented a large-signal table-based FET model. The model is capable of accurate small-signal modeling over the wide bias-range and efficient large-signal power modeling. The model has been validated by comparing the small-signal s-parameters and power performance with the measured data. The model extrapolation beyond the measurement range was taken into account.

REFERENCE

1. D. E. Root, S. Fan, and J. Meyer, "Technology-independent Large-signal FET Model: A measurement-based Approach to Active Device Modeling," 1991, *Proc. 15th ARMMS Conference*
2. A. Werthof and G. Kompa, "A unified Consistent DC to RF Large-signal FET Model Covering Strong Dispersion Effects of HEMT Devices," 1992 *EuMC*, pp. 1091-1096
3. M. C. Filicori and G. Vannini, "Mathematical Approach to Large-signal Modeling of Electronic Devices," 1991, *Electronic Devices*, Vol.27, pp.357-359
4. I. Schmale, F. van Raay, and G. Kompa, "Dispersive Table-based Large-signal FET Model Validated in Analysis of MMIC Frequency Doubler," 1996, *Proc. 26th EuMc*, pp.260-263
5. C. J. Wei, J. W. Bao and J. C. M. Hwang, "Novel Experimentally-based Modeling for MESFETs Having Complex Dispersion Effects," 1997, *IEEE Workshop on Experimentally-based FET Device Modeling & Related Nonlinear Circuit Design*, pp.22.1-22.8
6. C. J. Wei, Y. A. Tkachenko and D. Bartle, "A Compact Large-signal Model of GaAs MESFET and It's Parameter Extraction," 1997, *Microwave J.* No.12

Electrical Package Modeling for High Power RF Semiconductor Devices

Tao Liang, Jaime Plá, Mali Mahalingam
Motorola Wireless Infrastructure Systems Division
5005 E. McDowell Road, MD E108
Phoenix, Arizona 85008
Tel: (602) 244-7677 Fax: (602) 244-5898
Email: R39001@email.sps.mot.com

Abstract

A modeling procedure is developed for generating electrical package models for air-cavity metalized ceramic packages. Such models are capable of accounting for parasitic effects associated with the capacitance of the bonding pads in the package and the inductance of the bond wires. We have applied this modeling procedure to extract package equivalent circuit models for some of Motorola's LDMOS RF power transistors. The models have been validated up to the third harmonic with respect to the device's operating frequency. It has been demonstrated that the package parasitic effects contribute significantly to an accurate prediction of the packaged transistor nonlinear performance.

Introduction

At RF and microwave frequencies, the package electrical parasitics need to be correctly modeled in order to predict the performance of packaged devices. For example, RF and microwave transistors are typically designed to operate at certain desired conditions, such as tuned for power and/or efficiency. The matching circuits therefore have to be built around the transistors. At high frequencies, the inductive effect may come from the bondwires which are necessary to provide electrical connection from transistor die to the outside circuit. If this effect is not properly modeled and included in the circuit simulation, the impedance condition will be incorrect. The package parasitics can sometimes be utilized to a certain advantage. For example, a transistor may include an internal matching circuit inside the package to increase the usable bandwidth, where the bondwire parasitic inductance can be used as a circuit element. In either case, the package model will be important for correctly predicting the circuit performance.

Our objective is to derive equivalent circuit models for the package parasitic effects which can be used in the overall packaged transistor models. This paper will present some of our original package electrical modeling efforts for the ceramic air cavity packages used for Motorola's LDMOS RF power transistors.

Package Model Extraction Procedure

There are typically multiple active cells in a high power RF device, which will contain many bond wires to make electrical connections from the active device to the package leadframe, as illustrated in Fig. 1. Under this situation, the package inductive effects arise from the self and mutual inductance of the bond wires, and the capacitive effects from the package bond pads. We can represent the package by the input package model and the output package model cascaded with the active device, as illustrated in Fig. 2. Notice that the package has one lead for the input and another lead for the output. Each package half consists of a shunt capacitance to ground, C_{pad} , and a series inductance, L_g for gate side and L_d for drain side.

To increase device output power, more active cells may be used in parallel. Doing so will also reduce the input impedance and device usable bandwidth. To improve the device performance, an internal matching circuit can be built into the package. For example, a single tee-section (series L -shunt C - series L) input matching network can be incorporated into the package, as illustrated in Fig. 3. This type of package can be represented by an equivalent circuit shown in Fig. 4, where the wirebonds form the series inductors L_{g1} and L_{g2} and a SiO_2 MOS capacitor is the shunt element C_{moscap} .

To obtain the equivalent circuit model parameters, we used a 3-dimensional electromagnetic simulator (e.g. Ansoft HFSSTM) and generated the S-parameters of the structure. Then we used a circuit simulator (e.g. HP-EEsof LibraTM) to extract the equivalent circuit element values. The extraction involved multiple steps. We first obtained S-parameters for the empty package, from which the equivalent capacitance, C_{pad} , of the bonding pad for both gate and drain side window frame of the package was extracted. We then obtained S-parameters for the package with gate or drain side bondwires directly bonded to a grounded metal. The pedestal had the same height, length, and the location as the real active device so that the bond wires kept the same geometric shape. The bondwire inductances of the gate L_g or the drain L_d were extracted with the same C_{pad} (fixed value, i.e. not optimizable) obtained from the empty package simulation.

In the package models, we neglected the coupling between the gate side and drain side of the package which may include the capacitive coupling between the package leadframes and mutual inductive coupling between the gate and drain wires. We had performed the simulations and found that such coupling is insignificant over the frequency of interest when compared to active device reverse isolation (S_{12}). The bondwires are typically made of Au or Al. The loss associated with finite conductivity of the metal could be included as a resistance in series with each inductance. We found that the resistance is typically in the order of 0.05 Ohm and can be neglected. The package model may take a form of three port network to account for any inductance to ground, which will depend on the definition of ground plane.

Modeling Results

We have applied this modeling method to the NI μ 250 package (case 360B-01), an air cavity metalized ceramic package. The package construction has CuW flange with alumina window frame with an air cavity measuring approximately 260 mils x 110 mils x 20 mils. The extracted bond pad capacitance, C_{pad} , was found to be 2.2 pF. Since the active transistor is offset in the package, the gate and drain wires are not symmetrical. We looked at different bond wire configurations, i.e., different number of bond wires for the gate and drain side (1.5 mil diameter Au). The S-parameters of the package can be obtained either by electromagnetic simulation or by directly measuring the package with bond wires using a network analyzer. The S-parameter data is then used by a circuit simulator to extract equivalent circuit values. Fig. 5 shows the results for the gate side inductance, in which the cross marker denotes the inductance obtained from the simulated S-parameters and the solid marker denotes the inductance obtained from the measured S-parameters. By curve fitting the simulated data, we obtained an analytical expression of the wire inductance as a function of the number of bond wires. This expression is useful for interpolating the inductance values for different number of wires. Similar results are observed for the drain side wires as shown in Fig. 6. We see good agreement between the simulation and measurement approaches.

To examine the package electrical model and its role in predicting the overall performance of the packaged transistor, we extracted the package model for Motorola's MRF183 LDMOS RF transistor. The MRF183 is composed of 24 active cells (121mm gate periphery) and uses an NI μ 250 package containing 12 gate wires and 13 drain wires. Its output power is rated at 45 Watts at 1 GHz. Its electrical performance at 945 MHz was measured with the use of passive input and output matching networks. Fig. 7, 8, and 9 show the measured power gain, drain current, and drain efficiency, respectively, as a function of the input power on several packaged transistors (shown as solid lines) biased at $V_{dd}=28$ V and $I_{dq}=250$ mA. We used the method described above to obtain the package model resulting with

values of $L_g=0.52$ nH, $L_d=0.36$ nH, and $C_{pad}=2.2$ pF. This package model is used in the circuit simulator, with the active device model (Root model), as well as models for the matching networks used in the measurement. The simulation results (shown as triangular markers) match the measured results very well. If the package model is not incorporated, the simulation with active device model and the matching networks only (shown as cross markers) predicted a lower output power and drain current due to incorrect impedance predictions. Similar behaviors were observed on the measured vs. simulated power gain and drain efficiency. This indicates the significance of properly accounting for package parasitic effects.

Conclusion

We have developed a modeling procedure which can be used to extract package equivalent circuit models for Motorola's LDMOS RF power transistors in metalized ceramic packages. The models have been validated up to the third harmonic with respect to the device's operating frequency. It has been demonstrated that the package parasitic effects contribute significantly to an accurate prediction of the packaged transistor nonlinear performance.

Acknowledgment

The authors would like to thank Mike Cardullo and Garry Funk for model validation data and Jian Peng and Peter Aaen for technical discussions.

Ansoft HFSS is a trademark of Ansoft
Libra is a trademark of HP-EEsof.

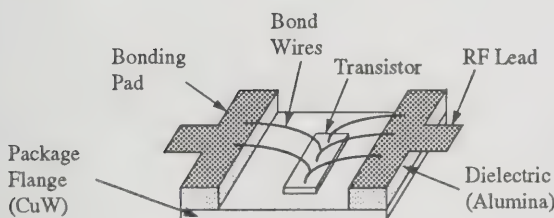


Fig.1 Typical package for RF power devices.

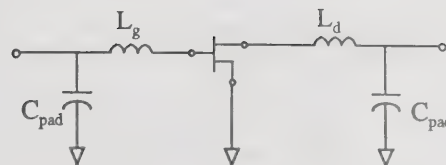


Fig.2 Equivalent circuit for a packaged transistor.

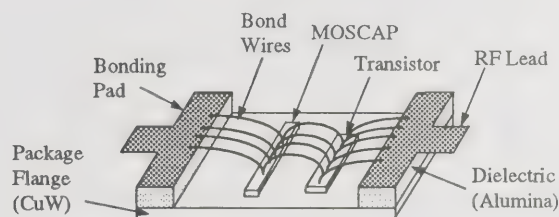


Fig.3 Package with internal matching capacitor.

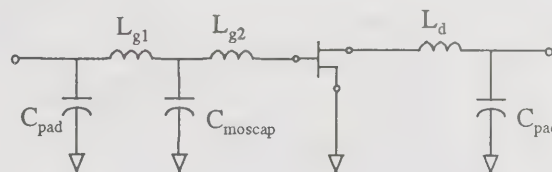


Fig.4 Equivalent circuit for the packaged transistor with internal matching capacitor.

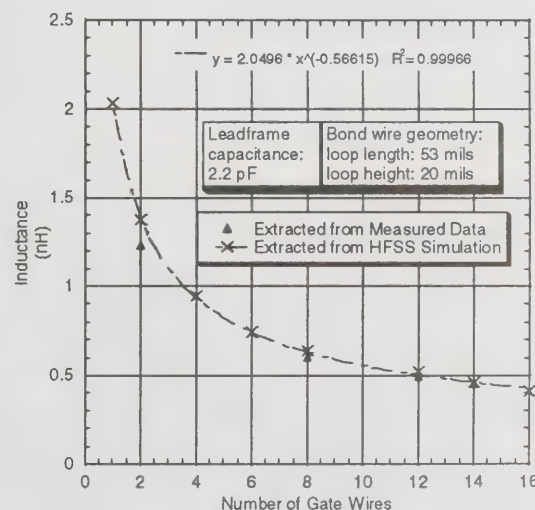


Fig.5 Gate wire inductance of NIμ250 package as a function of number of wires.

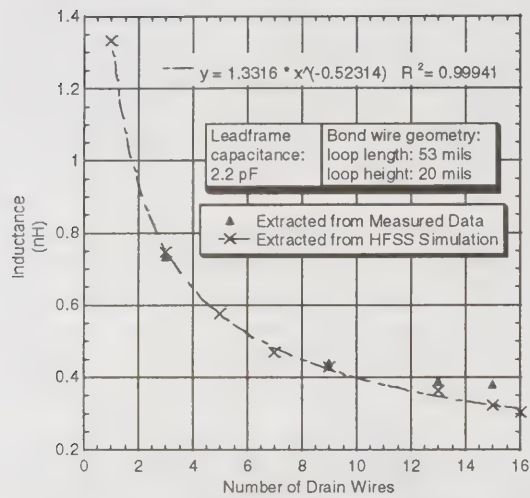


Fig.6 Drain wire inductance of NI μ 250 package as a function of number of wires.

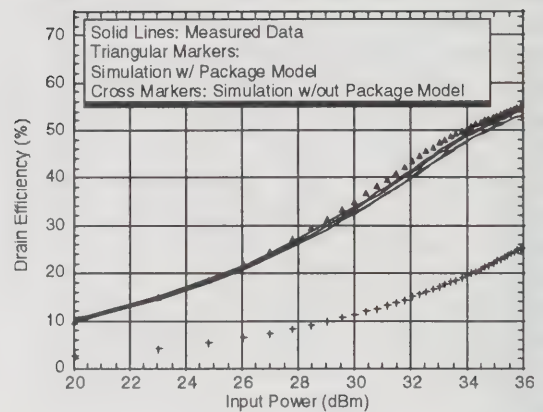


Fig.9 Measured and simulated drain efficiency.

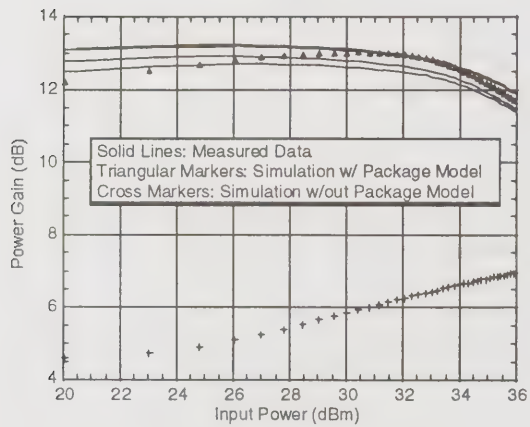


Fig.7 Measured and simulated power gain.

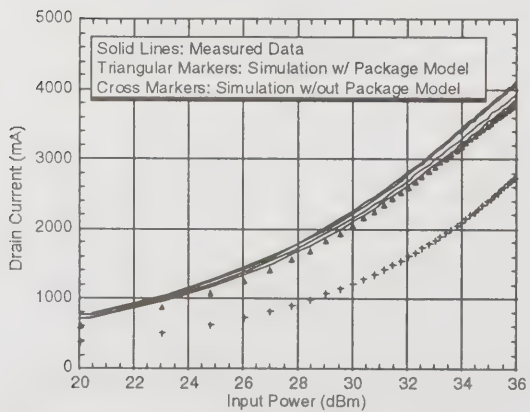


Fig.8 Measured and simulated drain current.

Index of Authors

Abraham, S.	43	Gracia, C.	43	Loyka, S. L.	233	Rey, C.	325
Achatz, R. J.	257	Greenberg, M. C.	87	Lu, W.	99	Rinaudo, S.	347
Ahmad, A.	1	Gregoire, D.	5	MacDonald, N. C.	269	Rivera, B.	59
Ainspan, H. A.	133	Gruenbacher, D. M.	63	Maeng, S.-J.	121	Ro, C. W.	95
Akaike, M.	249	Guerreiro, A. M.	209	Mahalingam, M.	359	Rohmer, G.	39
Al-Dabbagh, A.	245	Gusmao, A.	15	Makelainen, T.	79	Rolando, A.	173
Alspector, J.	165	Gutmann, R.	285	Maloney, P. R.	317	Rosenberger, B.	103
Ambardar, S.	31	Han, S.-M.	273	Mandyam, G.	193	Rudokas, R.	153
Arcioni, P.	301	Hanke, G.	113	Matsumura, K.	47	Sacchi, E.	301
Bächtold, W.	125	Harsh, K. F.	293	Mayaram, K.	351	Sadowiczak, G.	313
Bartle, D.	355	Hassaine, N.	281	Mayorgas, L. F.	261	Schaffner, J.	5
Bergstrom, C.	27, 169	Hazouard, M.	339	McGeehan, J. P.	137	Scholtz, R.	67
Beukema, T. J.	133	Helms, D.	329	McLean, J.	217	Scott, I.	35
Boehm, K.	39	Hentschel, T.	39	Mikuteit, S.	217	Shen, Y.	281
Bolton, D.	289	Ho, H. C.	265	Milstein, L. R.	197	Sherman, R.	313
Bos, T. A.	125	Hofstetter, P. M.	145	Mohan, A.	253	Shirakawa, I.	47
Brakensiek, W.	309	Hsu, H. P.	5	Mordachev, V. I.	233	Sievenpiper, D.	221
Brinegar, C.	277	Hu, C.	351	Mueller, T.	39	Smithgall, D.	157
Brown, T. X.	51	Hummels, D. R.	63	Muraguchi, M.	129	Soyuer, M.	133
Burghartz, J. N.	133	Humphrey, M.	59	Na, N.	117	Staudinger, J.	313, 343
Canora, F. J.	133	Inada, H.	47	Naishadham, K.	277	Strangeways, H. J.	237
Castello, R.	301	Irwin, R. S.	293	Neto, A. D.	209	Suoranta, R.	79
Chan, K. T.	99	Itoh, T.	221	Neves, H. P.	269	Susman, L.	109
Cheng, Y.	91	Jackson, D. L.	75	Ng, C. H.	205	Suzuki, H.	253
Choi, K. L.	117	Jarry, P.	339	Nohr, W.-D.	113	Svelto, F.	301
Chrisikos, G.	141	Jarvinen, E.	321	Norgard, J. D.	165	Swaminathan, M.	117
Chuprun, J. S.	27, 169	Jennings, D. J.	137	Norris, G.	313	Testa, M.	329
Clark, C. J.	141	Jimenez, F. A.	261	Ntake, P.	281	Todd, C.	59
Clark, E.	325	Kerherve, E.	339	Oehler, F.	39	Tin, S. F.	351
Clearfield, H. M.	289	Kim, K.	1	O'Farrell, T.	23, 177	Tkachenkov, Y. A.	355
Colella, N.	9	Kim, S.	83	Omiyi, P. E.	23	Upton, D.	317
Couch, L. W.	83	Kim, S. G.	95	Oprysko, M. M.	133	Uysal, S.	205
Crook, G.	217	Kim, Y.-G.	121	Osman, A. A.	351	Vaitkus, R.	343
Dalke, R.	257	Kim, Y.-S.	273	Parish, J.	165	Van Lil, E.	161
Darnell, M.	245	Koert, P.	109	Park, C.-S.	121	Vangala, R.	297
Day, P.	75	Kormanyos, B. K.	333	Park, S. M.	95	Virga, K. L.	87
de Baca, M. C.	343	Krishnamachari, B.	43	Patel, N. R.	177	Vladimirsky, N.	329
De Coster, I.	161	Kruse, R. W.	333	Paver, N. C.	75	Vojcic, B. R.	197
Diamandis, P.	9	Kudrle, T. D.	269	Payne, J.	185	Wang, S.-X.	241
Dinis, R.	15	Kuhn, W. B.	305	Perez-Fontan, F.	161	Weber, D.-J.	113
El-Hefnawi, F. M.	201	Lau, J.	265	Perregrini, L.	301	Wei, C. J.	355
Eriksson, T.	79	Lazear, W.	31	Pham, J. M.	339	Welch, T. B.	189
Ettus, M.	55	Lee, J.	273	Pham, T. D.	229	Wickert, M. A.	145
Featherstone, W.	237	Lee, J.-H.	121	Pilmanis, E.	133	Wills, D.	329
Ferla, G.	347	Lee, J. M.	95	Piolini, F.	173	Wong, T.	71
Fernandes, C.	225	Lee, Y. C.	293	Pla, J.	359	Wood, A.	309
Fernandes, J.	225	Leong, M. S.	205	Plouchart, J.-O.	133	Yablonovitch, E.	221
Fernandes, J.	149	Leonov, V.	109	Poon, M. C.	265	Yamagishi, A.	129
Flores, S. J.	261	Leuven, K.	161	Posti, J.	79	Yanduru, N. K.	305
Follensbee, D. R.	333	Liang, T.	359	Privitera, G.	347	Yang, S.	181
Fujita, G.	47	Lien, W. A.	75	Proakis, J. G.	197	Yeung, T.	265
Izadpanah, H.	5	Lisboa, F. A.	209	Qian, Y.	221	Young, J. L.	289
Galluzzo, A.	347	Liu, J.	75	Qin, T.-K.	91	Zhang, W.	293
Gao, S.	213	Logan, E. A.	289	Qureshi, A.	185	Zhang, X.	91
Gaucher, B. P.	133	Lok, S.	43	Radisic, V.	221	Zhang, Y.	241
Gavilanes-Loureiro, B.	161	London, S.	109	Ramirez-Mireles, F.	67	Zhong, S.	213
Glibbery, R.	75	Lott, U.	125, 335	Reudink, M.	19	Ziemer, R. E.	189, 197



1998 IEEE Radio and Wireless Conference

Sheraton Colorado Springs

Colorado Springs, Colorado, USA

August 9-12, 1998

<http://rawcon.org>
